MAR 19 1929

PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS

HAROLD RUGG, *Chairman.*

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN

FRANK N. FREEMAN — V. A. C. HENMON — BEARDSLEY RUML — E. L. THORNDIKE

VOL. XX
No. 3

MARCH, 1929

$5.00 a year.
70c per copy.

CONTENTS

THE CHILD'S EARLY EXPERIENCE WITH THE LETTER *A*. *Ernest Horn*. 161

A COMPARISON OF TWO TYPES OF SILENT READING AS USED BY CHILDREN IN DIFFERENT SCHOOL GRADES. *C. T. Gray* . . 169

THE RESULTS OF RESCORING FIVE HUNDRED THIRTY DEARBORN TESTS. *Walter F. Dearborn and C. Wilson Smith* 177

PHOTOGRAPHIC MEASURES OF READING ABILITY. *Miles A. Tinker* 184

FIRST IMPRESSION VERSUS SECOND THOUGHT IN TRUE-FALSE TESTS. *Mary Louise Lowe and C. C. Crawford* 192

THE HIGH SCHOOL STUDENT AND SCIENTIFIC METHOD. *M. Louise Nichols*. 196

SPELLING BY RULE. *Cornelia R. Trowbridge*. 205

THE UNIQUENESS OF "G." *C. Spearman*. 212

COMMENT ON PROFESSOR SPEARMAN'S NOTE. *Edwin B. Wilson* . 217

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES. 224

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION. 231

Published Monthly Except June to August by

Warwick & York, Inc.,

10 East Centre Street — Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at Baltimore, Md., under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Rugg, Chairman of the Editorial Board, 400 W. 118th Street, New York City.

The price of the Journal is $5.00 a year, payable in advance. Foreign postage is 30 cents extra. Single current issues are 70 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered. Back volumes $5.50 each; single back issues 90 cents each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for nonreceipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,
Baltimore, Md.

TEACHERS' HEADQUARTERS

European
Plan

Entirely
Fire-Proof

Centrally
Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

The Value of the Dictionary

is the title of a *new* booklet written to help you in Teaching the Dictionary. Here are a few suggestions of the lessons included:

First Dictionary Lessons
Relative Position of Letters
How to Find Words
What You Find
Pronunciation
How to Find Meanings
Parts of Speech and Meanings
Unusual Uses of Words
Synonyms
The Hyphen, Etc., Etc.

Copies of this new booklet will be sent FREE to teachers upon request.

G. & C. Merriam Company
Springfield, Mass.

Publishers of the Merriam-Webster Dictionaries for over 85 years
Look for the Circular Trade-Mark — Get the Best

WEBSTER'S NEW INTERNATIONAL DICTIONARY

SCHOOL AND SOCIETY

A weekly journal covering the field of education in relation to the problems of American democracy.

The objects of SCHOOL AND SOCIETY are the advancement of education as a science and the adjustment of our lower and higher schools to the needs of modern life.

It encourages discussion of all problems from every point of view; it pays special attention to foreign affairs; it publishes accounts of educational events of the past week and research on fundamental problems.

Subscription $5.00—Single Numbers 15 Cents

SCHOOL AND SOCIETY will be sent to new subscribers for thirteen weeks for one dollar, which may be sent at the risk of the publisher.

THE SCIENCE PRESS
Grand Central Terminal - - New York, N. Y.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XX March, 1929 Number 3

THE CHILD'S EARLY EXPERIENCE WITH THE LETTER *A*

ERNEST HORN

State University of Iowa

We have long been taught that learning is facilitated, and remembrance made more certain, by utilizing meaningful associations. It is not surprising, therefore, that we have with us always one or more varieties of plans for rationalizing spelling and reading. Learning to spell four thousand different words seems a large task if each word must be learned as an individual task. The total number of different words in the list of books which the child is expected to read runs far beyond this figure even in the first six grades. Learning to associate arbitrarily the appropriate meanings with each of these printed symbols seems a most formidable undertaking.

Neither in reading nor in spelling have attempts at rationalization given results which are very satisfactory. Children who have had training in phonics do not show, according to the rather meagre evidence which is available, any significant superiority in spelling. The investigations of teaching phonics in reading have been more extensive, though poorly controlled, but even here the investigators not only do not agree upon the type of phonic training that should be given; they also do not agree as to how much phonics, if any, should be taught. The failure to show the desirability of phonic training may be due to one or more of several possible causes. First, it must be admitted that the investigations of phonic *vs.* non-phonic training which have been made up to the present time have been poorly controlled. It is possible that were the factor of phonic training skillfully isolated and the results of the experiment adequately analyzed, a superiority would be shown. Second, it is probable that many children who have not been taught phonics have made their

own comparisons between words and have in that way built up a sort of phonic system of their own. It is a common thing to find some children who have had no phonic training but who show remarkable ability to pronounce words which they have never read before. Moreover, many so-called "intrinsic" methods really include phonic training. Third, an analysis of existing phonic systems shows that many of them either contain serious phonetic inaccuracies or are clumsily devised. Fourth, it is likely that few teachers teach any system of phonics as it is planned to be taught. Fifth, the unphonetic character of the English language may constitute an insurmountable barrier to successful rationalization. No doubt each of these factors helps to account for the fact that up to the present time the results of investigations are not in agreement as to the kind or amount of phonic teaching which should be recommended.

The following paragraphs present the results of a partial inquiry into the fifth possibility enumerated above; namely, that the English language is so unphonetic as to make impracticable attempts at rationalization through phonetic teaching. Let us begin with the first letter in the alphabet. What varieties of experiences may a child be expected to have with the letter *a* in Grades I to III? These are the grades in which most phonic instruction is given. Here, also, the problem of learning to read is most acute. Table I, which follows, attempts to give a summary view of the difficulties confronting the child in his attempt to rationally control this letter either in spelling or in reading. The data on spelling are presented on the assumption that he has studied throughout these grades a certain speller in wide current use. The part of the table which deals with spelling would not be greatly affected were some other speller chosen instead. The data on reading are based on the detailed tables in Dr. Cordts's[1] analysis of the representations of all of the sounds in ten first, ten second, and ten third grade readers. Low frequency words were also included from the sources used by Dr. Cordts.

---

[1] Cordts, Anna D.: "An Analysis and Classification of the Sounds of English Words in a Primary Reading Vocabulary." Ph.D. Thesis, unpublished, State University of Iowa, 1925.

NOTE.—While some of the books used in this investigation are no longer widely read in the primary grades, a sampling of modern books indicates that the results would not be greatly modified were this analysis made with an equal amount of modern primary reading material.

TABLE I.—THE CHILD'S EXPERIENCE WITH THE LETTER *A* IN GRADES I TO III

Sound	Example	Number of different words in spelling	Number of different words in reading	Total frequency in reading
(1) ā a	paper, April	9	41	2,177
a-e	ate, came	44	95	8,974
ai	fail, rain	9	55	2,769
ai-e	praise	1	6	156
ay	play, day	18	35	8,654
ea	great, break	1	5	1,015
(2) ĕ a	many, any	4	8	1,570
ai	said, again	2	3	6,979
ay	says	1	1	276
ea	bread, head	6	36	2,641
(3) â a	parent	0	1	14
a-e	bare, care	5	14	576
ai	air, hair	7	10	848
ay	prayer	0	1	11
ea	wear, bear	2	7	609
(4) ô a (wa, al)	warm, call	24	43	9,128
au	fault	1	6	554
augh	daughter, naughty	0	4	357
aw	paw, saw	5	19	1,629
oa	broad	0	1	32
(5) ŏ a	was, watch	7	16	9,096
(6) ē Ae	Aeolus	0	1	6
ea	eat	15	115	8,536
(7) ă a	at, and	62	216	57,443
a-e	have, bade	1	7	2,852
(8) ä a	ah, arm, car	20	80	5,744
aa	baa	0	1	25
a-e	are	2	3	3,058
au	laugh, aunt	1	4	427
ea (r)	heart, hearth	0	3	169
ua	guard	0	1	11
(9) å a	ask, pass	18	51	5,011
(10) ī ai	mountain, fountain	0	5	122
ai-e	marriage, carriage	0	2	20
(11) ā a	language	0	1	6
a-e	palace, village	0	4	274
ay	yesterday	6	8	111
(12) ȧ a	again, diamond	13	33	19,536
(13) ȁ a	distance, ocean	0	30	1,293
(14) ẽ a (wa, ar)	beggar, toward	1	15	429
oa	cupboard	0	1	31
(15) û ea (r)	search, early	3	12	1,095
(16) ō oa	boat, board	4	31	1,504
eau	bureau	0	1	4
(17) ū eau	beauty	0	2	581
(18) ɪ ay	ay	0	1	1
a-e	aye	0	1	1

All pronunciations and phonetic transcriptions in this and in other tables are based on Webster's "New International Dictionary." This standard was chosen, first, because it is well-known to teachers and pupils, and second, because the writer's limited training in phonetics made it impossible for him to bring order out of the confusion arising out of the fact that authorities in phonetics disagree not only in regard to the appropriate transcriptions of the pronunciation of many of the words but also with regard to the pronunciation itself.

This table should be read as follows. The child will see the sound ā as in *paper* made in six different ways by utilizing the letter *a*: a, paper; a-e, ate; ai, fail; ay, day; and ea, great. (There are, of course, other ways of making this sound without utilizing the letter *a*.) He will spell forty-four words in which the sound is made by using *a* followed by a consonant and final *e*. He will read ninety-five words with a total occurrence of 8974 in which the sound is made in the same fashion, etc.

The table includes forty-seven different sound-letter associations for the letter *a* in words actually occurring in First, Second, and Third readers. No one interpretation of the letter *a* has a majority of uses in its favor either as measured by the number of different words falling under it or the total frequency of occurrence as measured by running words. The most frequent value for *a* is the so-called short *a* (ă), which is found in sixty-three words in spelling, and two hundred twenty-three words in reading, with 60,295 occurrences in reading counting repetitions. But the tables show forty-five other types of sound-letter associations involving two hundred twenty-nine words in spelling and eight hundred thirteen words in reading with 106,060 occurrences counting repetitions. It should be noted that the child's experience with most of these forms, as measured by the total frequency of occurrence in reading, runs into the hundreds and in nearly half of the situations into the thousands. In the light of repetitions usually considered ample for drill and review purposes, these frequencies loom very large.

The following comments, numbered to correspond to the divisions in the table, may help to make the table more intelligible.

(1) It will be seen that no one of the seven letter groups which express the letter *a* includes as much as half of the total cases, either as measured by the number of different words or by the total frequency in reading. There are, of course, other ways to make this sound, as for example, by using *au*, as in *gauge*, as well as by using letters, or groups

of letters, which do not include the letter *a* at all, as in *they*. As will be seen by examining other parts of the table, every one of the letter groups which here express ā also expressed at least one other sound.

(2) It is possible that this group should have had as an additional classification a-e to include such words as *orange*. Such words were classified under *a* in this section of the table. The sound ĕ is also made with *ae* as in *aesthetic*, but no such words occurred in the readers for the first three grades.

(3) The sound â is given a separate classification in the table to agree with the usage in Webster's "New International Dictionary." It is probable that a good many children would not distinguish the sound of *ai* in *air* from the sound of *ai* in *said*.

(4) It is possible that this division of the table should have had also the letter grouping au-e as in *because*. Such words were classified under *au*.

(5) It is possible that the single classification under (5) should have been divided into three parts, the first to include words beginning *wha*, the second, words beginning *wa*, and the third, words which contain one or more syllables ending in *ar*, as in *war*.

(6) Under *ea* in this table are included words like *eat* as well as words like *hear* and *ear*, since these words are given the same transcription in the "New International Dictionary." Many authorities in phonetics do not attach the same value to the *ea* in these words. The word *Aeolus* is, of course, a joker, but it actually occurred in one reader. The word *Aesop* would not seem so much out-of-place.

(7) It is interesting to note here a very early and frequent exception to the effect of the final *e* as given in most phonic systems.

(8) *Guard* is placed in this list since the *ua* in this word is listed in the "New International Dictionary" as a digraph. *U* after *g* usually has the effect of indicating that the *g* is hard.

(9) In this group words like *dance* are classified under the letter *a*, although a separate division might conceivably have been made under a-e, even though the *e* is preceded by two consonants.

(10) A separate group is made in (10) and (11) for the sounds ĭ and ä̇ according to the usage in the "New International Dictionary," even though these sounds shade into each other in cultivated speech.

(12) It is interesting to note in the word *diamond* that the *ia*, which appears several places in the table as a digraph, is here separated into the vowels *i* and *a*, which appear in different syllables.

(13) It is interesting to note that in *ocean* the *e*, which commonly unites with *a* to form a digraph, here is separated from *a*, and with *ce* makes the sound *sh*.

(18) The sound ī is also made with *ai* as in the word *aisle*, but although the child probably knows this word in the first three grades, it was not found in any of the readers used by Dr. Cordts. The words *ay* and *aye* are here counted as meaning *yes* and included under the sound ī, although it is possible that in the readers the words were used with different meanings, in which case they should appear under ā.

The confusion pictured in Table I is much worse than it would be if the child were confronted with forty-seven wholly different types of situations involving *a*, since here the same letter combinations cannot be consistently interpreted. Table II illustrates this difficulty by showing the variety of ways in which the child must interpret the digraph *ea*.

TABLE II.—HOW SHALL THE CHILD INTERPRET *ea*?

	Number of different words in spelling	Number of different words in reading	Total frequency in reading
ā great	1	5	1,015
ĕ bread	6	36	2,641
â wear	2	7	609
ē eat	15	115	8,536
ä heart	0	3	169
û search	3	12	1,095
ū beauty	0	2	581
ō bureau	0	1	4
Total.....	27	181	14,650

This table should be read: *ea* is sounded ā in one word in spelling and in 5 different words in reading with a total frequency, counting running words, of 1015, etc. While the most frequent sounds expressed by these combined letters is ē, more than a third of the occurrences as measured by running words are pronounced differently. There are eight different sound values for *ea*. These various sounds for *ea* are scattered through his reading with so little plan as to approach a chance distribution and must be a source of great confusion.

The other letter combinations follow the same sort of variation in sound values. Thus, *a* expresses twelve sounds; *ai* expresses four

sounds; *ay*, five; and even *a* followed by a consonant and so-called silent *e* expresses six sounds. Today the child must sound *oa* as in *broad*, tomorrow *oa* as in *boat*. Moreover, the child will soon be confronted by new values for *a*, as ī (*aisle*) and ā (*gauge*).

Indeed, from the point of view of the primary child there are additional complications due to the fact that some of the vowels with which the letter *a* sometimes forms a digraph are not infrequently separated from *a* in syllabication, influence the sound of consonants, or are actually used as consonants. Illustrations of these difficulties are given in Tables III and IV.

TABLE III.—EXAMPLES OF WORDS IN WHICH *a* IS ADJACENT TO ANOTHER VOWEL WITH WHICH IT COMMONLY FORMS A DIGRAPH BUT FROM WHICH IT IS IN THESE INSTANCES SEPARATED IN SYLLABICATION

	WORDS IN PRIMARY READERS USED BY CORDTS	OTHER WORDS
ea	idea, Indian	
ai		naive, archaic, Isaiah
ia	diamond, piano, giant	diameter, liar
ae		aerial
oa	Noah	boa
ua	truant, gradual, January, February	
ao		chaotic, aorta

A real puzzle is furnished the child by the word *hyacinth*, which appears in one of the readers. Moreover, he will soon meet extraordinary (ĕks-trôr′dĭ-nȃ-rĭ, ĕks-tra-ôr′dĭ-nȃ-rĭ).

TABLE IV.—EXAMPLES OF WORDS IN WHICH *a* IS ADJACENT TO A VOWEL WITH WHICH IT COMMONLY FORMS A DIGRAPH BUT WHICH IN THESE INSTANCES IS USED TO INFLUENCE THE SOUND OF A CONSONANT, IS USED AS A CONSONANT, OR HELPS TO MAKE A CONSONANT AS IN *qu* AND *ti*

	WORDS IN PRIMARY READERS USED BY CORDTS	OTHER WORDS
(g) (s) ua	language, guard	persuade
(c) (t) ia	physician, musician, Christian	filial, partial, initial, genial
(c) ea	ocean	
q (ua)	quality, quack	

The actual difficulty confronting the child in interpreting a new word is not quite so great as the tables indicate since approximate sounds, together with the reading context, may give a sufficient clue to the word, but even when the most liberal allowance is made for such approximations in reading, the confusion here pictured is still serious. In spelling, where approximations are not acceptable, the confusion is, of course, much greater.

One cannot, of course, conclude from these data that phonics should not be taught. One can only conclude that plans for teaching phonics must take into account such facts as are here presented. When this is done, the results of teaching phonics will undoubtedly be more satisfactory. For while the unphonetic character of the English language constitutes a real obstacle to successful rationalization, it is not necessarily an insurmountable one. But the degree to which this obstacle may be surmounted by phonic training, what kind of phonic training is best, and how or when it should be given—these are questions for which, at present, there are no answers upon which investigators in reading agree.

A COMPARISON OF TWO TYPES OF SILENT READING AS USED BY CHILDREN IN DIFFERENT SCHOOL GRADES

C. T. GRAY

University of Texas

The problem of the investigation at hand is to determine the extent to which children in Grades IV, V, and VI are able to use two different methods or techniques of silent reading. The two methods selected for comparison are (1) reading for the purpose of answering questions when the questions are given before the reading is done, and (2) reading for the purpose of answering questions when the questions are given after the reading is done.

As the problem is treated, it does not raise the question as to which is the better method for children to use, but has to do only with the pupil's ability to make effective use of two types of reading which are similar in many respects and fundamentally different in other ways.

In carrying out the problem for this standpoint, it is helpful to have in mind the nature of the reading process of experienced readers when they are called upon to make use of the two methods of reading. Adult introspection indicates that the first type of reading has in it three distinct elements or phases. In this type, to be referred to as Method I, it is advantageous first to decide upon the key word or words in the question for which the answer is to be found. Second, it is advisable to look through or skim the reading material until words or sentences are encountered which seem to relate to the key word in the question. Third, it is necessary to read in detail that part of the selection where the answer to the question is found.

With respect to the second type of reading, to be referred to as Method II, introspection indicates: First, that this method must be accurate and detailed because the questions to be answered may concern any or all parts of the selection. Second, success in this method of reading depends in some degree upon the anticipation of the questions to be encountered later. Third, such reading demands that details of the selection as read be held in memory for at least a short period of time.

If the above analysis is correct, the first method can be described as an immediate, somewhat gross process which depends upon a considerable degree of fluency, while the second method involves a more remote end and more detail.

Experimental work with adults has shown that in the case of the well-trained and experienced reader, use of the two methods yields different results. Method I allows a greater rate of reading because a part of the materials has only to be skimmed. Method I also allows more accuracy in interpretation because memory as a factor is eliminated. Furthermore, the correlation between the two methods is high because the experienced reader obtains success in the use of each method.

Five studies typical of work on this general problem will be cited. Ritter and Lofland[1] have shown that correlation between scores based upon a so-called general reading ability and reading ability where reasoning ability is involved in a large degree, is less in the upper grades than in the lower grades. This indicates that specialized reading is acquired in the upper grades.

Judd and Buswell[2] studied the problem of different methods of reading by means of the detailed technique of eye-movements and found that as the purpose of reading changes, the method of reading changes.

Good[3] has shown that among pupils of high school age mental set is an important element in their reading. Gates'[4] study of various reading tests is also in point at this place.

A different approach to the same problem is being made by those who are immediately responsible for instruction in intermediate and upper grade reading. Courses of study covering this period of school work give lists of different kinds of training which children are to receive in reading, as reading for the purpose of outlining, answering questions, etc. In this connection Neal and Foster[5] found from a study of textbooks for the fourth and fifth grades that fifteen types of

---

[1] Ritter, B. F., and Lofland, W. T.: The Relation between Ability in Reading General Material and in Reading Material Involving Reasoning. *Elementary School Journal*, Vol. XXIV, p. 529.

[2] Judd, C. H., and Buswell, G. T.: "Silent Reading; A Study of the Various Types." Supplementary Educational Monograph No. 23, Nov., 1922, University of Chicago Press.

[3] Good, Carter V.: The Effect of Mental-set or Attitude on the Reading Performance of High School Pupils. *Journal of Educational Research*, Vol. XIV, p. 178.

[4] Gates, A. I.: An Experimental and Statistical Study of Reading and Reading Tests. *Journal of Educational Psychology*, Vol. XII, pp. 303, 378, 445.

[5] Neal, Elmer A., and Foster, Inez: A Program of Silent Reading. *Elementary School Journal*, Vol. XXVII, p. 275.

reading are required of children in these grades. The same authors have made a careful analysis of reading for different purposes and give sample lessons of various types.

As indicated in the study of Neal and Foster a curriculum based upon specialized types of reading for intermediate and upper grade reading is slowly evolving. The importance of the problem at hand lies in the fact that it attempts to answer the question as to whether children need such training, since some have argued that a single type of general training is all that is required.

The results to be reported were obtained from three school systems,[1] Ruston, La., San Antonio, Texas and Austin, Texas. A total of four hundred ninety-one children were used in the experiment. Since space is limited and since the three groups of data cannot be combined, results from only one of these systems will be given in detail.

As a basis for this work, two standard silent reading tests were used, Monroe's test and the Thorndike-McCall test. In the Ruston and San Antonio schools both of these tests were used. In the reading according to Method I the children were told to look at the questions first, while for Method II the test was given in the usual way.

In the Austin schools, only the Thorndike-McCall test was used. For this work permission was obtained from the authors to reprint Form 1 of this test so that the questions in each case appeared above the selection to which they related. It was thought that the test in this form would aid the children in carrying out the reading according to Method I. Form 1 as modified and Form 3 without change were given in Grades IV, V and VI of two schools. Form 3 was given first according to the standardized directions except that in order to get a rough measure of speed the time was reduced to eight minutes in each of Grades V. One week later Form 1 was given. The directions to the children for this form necessarily had to be changed slightly.

Since there is no way to tell what changes the revised form of this test produced in the relative difficulty of the passages used in the test, it was thought best to use the number of questions answered correctly as a basis for scoring rather than the T scores used by the authors. Results from these tests are shown in Table I.

---

[1] I am indebted to Miss Maud Harper, Louisiana Polytechnic Institute, Ruston, Mr. T. Guy Rogers, Principal of Collins School, San Antonio, and Supt. A. N. McCallum of Austin for help in bringing this material together.

TABLE I.—SHOWING RESULTS IN AUSTIN SCHOOLS

Arithmetic Mean for Number of Questions Answered Correctly

Grade	School 1			School 2		
	IV	V	VI	IV	V	VI
N	37	49	32	39	32	56
Method 1	17.3	15.2	23.5	19.6	16.3	24.2
Method 2	16.8	14.4	25.3	20.2	16.5	25.8
σ_{diff}	.87	.52	.72	.59	.59	.48

Range for Number of Questions Answered Correctly

Method 1	7–27	9–24	18–29	11–27	10–22	16–29
Method 2	2–26	7–23	17–34	7–26	11–22	18–30

The following points may be made in connection with this table: First the differences in the means for the different grades are small; second, the means show that two grades do better with the Method I than with Method II; third, the standard errors of the differences indicate reliable differences; fourth, the total range is greater for Method II in four of the six cases; and fifth, the upper limit of the range is greater for Method I in three cases and greater for Method II in two cases. When contrasted with results obtained from adult readers already mentioned these data indicate clearly that the pupils were not able to make accurate and effective use of the two methods.

Different factors may be operating to produce such results. First, it is possible that children of these grades are not able to appreciate the differences which exist in the two types of reading and therefore the reading done in the two tests is very much the same. Second, one of the two methods of reading may be inherently more difficult for some children than the other.

Results were next obtained by means of correlations. These data are shown in Table II. The following points are to be noted with respect to this table:

TABLE II.—SHOWING CORRELATION COEFFICIENTS BASED UPON SCORES FOR THE TWO METHODS OF READING

School 1			School 2		
Grade	*r*	PE	Grade	*r*	PE
VIB	.823	±.074	VIB	.735	±.073
VIA	.588	±.101	VIA	.456	±.113
VB	.568	±.110	VB	.230	±.184
VA	.775	±.047	VA	.647	±.089
IVB	.588	±.101	IVB	.426	±.179
IVA	.780	±.071	IVA	.775	±.046

First, the highest correlations are found in both schools in Grades IVA and VIB. Such results for children in Grade IVA may be due to the fact that at this stage of development children are unable to make distinctions with respect to the two types of reading. In other words, their reading is all of the same type, hence the correlation of the two forms of the test would be high for such readers. The results for grade VIB are probably due to the fact that children at this stage of their development are able, only in part, to make the distinctions called for in the two types of reading. As a result of this they do about as well in one test as the other. Such a situation would, of course, lead to a high correlation. The lower correlations in Grades IVB, VA, and VIA indicate that children in these grades are confused in some degree when confronted by the different methods of reading and as a result the correlations based upon the results of such reading are relatively low. The high correlation for Grade VA in School 1 is apparently produced by the fact that this grade happens to be made up almost entirely of poor readers. Their case becomes, then, like that of Grade IVA as cited above. The absence of uniformity indicates clearly the failure of the pupils to adjust themselves to the problem.

Another point to be stressed in connection with Table II is that all the correlations are lower than correlations found between different forms of the test by the author. Professor McCall reports to me through private correspondence that he has found a correlation of .80 between Forms 2 and 3, and that he thinks this is what may be found between other forms. If this statement of inter-correlation is taken as a kind of standard, then the fact that none of the correlations in Table II is as high as the standard indicates the operation of some

factor which brings about confusion in the reading activity of the children.

Certain results with respect to speed of reading will now be cited. It will be recalled that the time for the test in Grades VA and VB was limited to eight minutes. In estimating speed the words were counted in each question and in each paragraph. Each child was given credit for the total number of words up to and including the last question as well as the last paragraph involved. This is only a rough measure but the results as shown in Table III seem to have some value.

TABLE III.—SHOWS MEDIAN RATES OF READING IN WORDS PER MINUTE AND CORRELATIONS

School 1			School 2		
Grade	Method 1	Method 2	Grade	Method 1	Method 2
VB	64.87	53.40	VA	57.37	75.5
VA	53.32	63.25	VB	64.18	64.12
		Correlations			
VA	$r = .746 \pm .09$		VA	$r = .755 \pm .07$	
VB	$r = .344 \pm .02$		VB	$r = .717 \pm .08$	

The absence of uniformity in these results can be accounted for by saying that these children were unable to carry out the two different types of directions equally well.

The next step in the procedure was to find a single correlation coefficient for the combined results from each of the grades under consideration. These were as follows:

GRADE	r	PE
IV	.622	± .044
V	.632	± .044
VI	.651	± .044

The fact that the correlation of Grade VI is higher than either of the other grades, and that Grade V is slightly higher than Grade IV indicates that experience and maturity play some part in the two types of reading. The fact that the results for each grade approximate closely the results for the others indicates that whenever a considerable

number of strong and weak pupils are brought together in this way there is a balancing of the factors which determines ability in the two methods of reading.

The next step in the procedure was to find the correlation between the two methods when chronological age was made constant. These results are shown in Table IV.

TABLE IV.—SHOWING CORRELATION WITH AGE A CONSTANT FACTOR

AGE, YEARS	PE
9	.518 ± .034
10	.711 ± .042
11	.862 ± .021
12	.873 ± .021
13	.842 ± .041

The advantage of the data shown in Table IV over results in Table III is apparent. Not only is chronological age made constant, but intelligence is constant in a large degree. Since the nine year old group is in Grade IV, it is composed of accelerated children, but in spite of their degree of intelligence the correlation for this group is lower than for any other group. This indicates clearly the value of experience and training, because their training is at least one year less than that of any other group. Again, the ten year old group is making normal progress in school. This means that the intelligence of this group is not so high as that of the preceding group, but they have had more training by at least one year than the nine year old group. The higher correlation for this group indicates again the value of experience and training. The same line of reasoning will apply to the eleven and twelve year old group except that these groups contain some retarded pupils. If these retarded pupils were eliminated, it is reasonable to suppose that the correlation would be higher than it is.

That intelligence is a factor is shown in the thirteen year old group. All children in this group are retarded and in spite of their added experience and training the correlation for this group is lower than that of the other groups. These data, along with others not given here, show clearly the value of training for children who are called on to deal with the two types of reading under consideration.

The training along these lines as received by the children in the experiment has grown out of their daily work. It seems reasonable to suppose that specific training as a part of the work in silent reading would increase the efficiency of children in such work.

In conclusion three points should be brought out:

1. Children of Grades IV, V and VI are not able to use the two types of reading under consideration in the most effective way.

2. The regular work of the school in the various subjects increases efficiency in such work from year to year.

3. Apparently specific training would be a distinct aid in developing the two types of reading.

THE RESULTS OF RESCORING FIVE HUNDRED THIRTY DEARBORN TESTS

WALTER F. DEARBORN AND C. WILSON SMITH

Harvard University

In examining the successive mental ages of a group of boys of North European descent of Medford, Massachusetts, wide differences in the mental growth of the same individual over the period of four years, during which repeated measurements had been made, suggested an examination of the test papers from which the mental ages had been obtained. This examination, involving the rescoring of four hundred nine Dearborn A and of one hundred twenty-one Dearborn C Intelligence Tests, showed clearly that Professor Pintner's experiment in scoring[1] had revealed a situation not peculiar to the test he used.

The gross results of rescoring these tests are shown in Table I. Column 1 shows the number of tests in which the score remained unchanged; column 2, the number of tests changed by the rescore; column 3, the number of tests whose score was raised, with the number

TABLE I[1]

1	2	3	4
142	388	90 (1)	44 (1)
...	...	100 (2)	20 (2)
...	...	33 (3)	15 (3)
...	...	29 (4)	10 (4)
...	...	12 (5)	1 (5)
...	...	12 (6)	4 (6)
...	...	1 (7)	1 (7)
...	...	4 (8)	1 (10)
...	...	4 (9)	1 (12)
...	...	1 (11)	1 (15)
...	...	2 (14)	
...	...	1 (15)	
...	...	1 (17)	
142	388	290	98

[1] Both A and C tests are shown here. The Dearborn C has two month intervals in its mental age scale. This accounts for the larger number of cases in the second frequency under column 3.

[1] Pintner, Rudolph: Accuracy in Scoring Group Intelligence Tests. *The Journal of Educational Psychology*, Oct., 1926.

of months raised shown in brackets after the test frequency; column 4, the number of tests the score of which was lowered with the number of months lowered shown in brackets.

Seventy-three per cent of all tests contained error, with 6.4 per cent of all tests showing error of half a year or more. If only those tests whose score was changed be considered, 8.8 per cent of the errors were serious enough to change the mental age from six to seventeen months. Even more startling is the obvious introduction of constant error. Seventy-five per cent of all errors were those of underscoring. If the errors in scoring the Dearborn C Test be separated from the total errors, they show forty-seven per cent of all tests (121) underscored, or eighty-six per cent of all errors those of underscoring. In no single year was the error of underscoring less than the error of overscoring.

In 1922–23, seventy-three per cent of all errors were those of underscoring
In 1923–24, sixty-nine per cent of all errors were those of underscoring
In 1924–25, seventy-five per cent of all errors were those of underscoring
In 1925–26, eighty-six per cent of all errors were those of underscoring

Constant error of this sort, of course, changes the statistical appearance of the whole group. The median of the distribution of the Dearborn C mental ages was raised 1.6 months. This change is not great, but several assumptions of considerable importance have been made on changes not much greater than this. It must be remembered that this raising of the median was made despite the many and serious errors of overscoring, and does not accurately express the danger of inaccurate scoring. If one is considering individuals, and is considering their repeated measurements with a view to discovering something of their mental growth, inaccurate scoring may produce some interesting results. Let an error of fifteen months underscoring in 1922–23 coincide with an error of fifteen months overscoring in 1923–24, and a year of perfectly normal growth appears as something grotesque. But the desirability of protecting mental tests from greater inaccuracy than is inherent in their nature needs no lengthy dissertation for its proof. It is the way in which inaccuracy creeps into testing results after they reach the laboratory, and the way in which they may be prevented from creeping in that offers the true problem.

One is likely to jump to the conclusion that the most fruitful source of error in scoring will be found in items of a test scored subjectively—that there will be a difference of opinion between the person scoring and the person checking on the score. In this study, such difference of opinion was found. It was responsible for some of the "errors"

noted, but not for as many as one would have expected. An analysis of forty-one Dearborn A Tests (selected at random from the four hundred nine test papers) and of one hundred twenty-one Dearborn C Tests indicates surprisingly little variation in the average number of errors found in questions scored objectively and in those scored subjectively. Table II shows the gross results of rescoring the Dearborn C Tests, and Table III analyses those errors. In the first three items of the test, subjective estimate plays no part in the scoring. Item 1 consists of pictures which will "tell a story" if placed in the proper

TABLE II.—ERRORS IN SCORING ONE HUNDRED TWENTY-ONE DEARBORN C TESTS

I	II	III	IV
55	66	39 (2)	7 (2)
..	..	12 (4)	2 (4)
..	..	5 (6)	
..	..	1 (8)	
—	—	—	—
Totals...55	66	57	9

Column I—number of tests unchanged in score by check.

Column II—number of tests changed in score by check.

Column III—number of tests raised in score by check with mental months raised shown in brackets.

Column IV—number of tests lowered in score by check with mental months lowered shown in brackets.

TABLE III.—ANALYSIS OF ERRORS IN SCORING FOUND IN ONE HUNDRED TWENTY-ONE DEARBORN C TESTS

I	II	III	IV
1	108	9 (1)	2 (1)
..	...	2 (2)	
2	192	12 (1)	4 (1)
..	...	1 (2)	1 (2)
..	...	1 (3)	
3	118	3 (1)	
4	70	31 (1)	8 (1)
..	...	9 (2)	1 (2)
..	...	2 (3)	

Column I—the number of the item of the test.

Column II—the number of tests unchanged by check.

Column III—the number of tests raised by check with the mental months raised shown in brackets.

Column IV—the number of tests lowered by the check with the mental months lowered shown in brackets.

sequence, they are numbered in order by the testee, and the number placed by him on each picture is either right or wrong. Item 2 is a series of words numbered in the same way and scored as either right or wrong. Item 3, the Block Test, has three things for the scorer to look for—the formation of a triangle made by drawing a single line between other lines, the number placed by the testee in the triangle so formed, and the printed number or numbers in each figure—which may not be used again by the testee. Here again, there is no possibility of subjective estimate except in the drawing of the line to form the triangle, and this is so slight as to be negligible. The fourth item of the test, an opposite test, appears to be as purely objective as do the other three; in reality, it has a decided subjective element. The scoring directions give a list of words which are to be considered as the only right answers, any departure from which should be scored "wrong" by the scorer. It would seem easy to score this test without error, certainly without error of a subjective nature. Actually, it is not easy for a person interested in his work to mark as wrong those answers which depart from the strict letter of the law but which he feels show an understanding of the question. The difficulty here lies in the richness of the English language, which makes it possible to repeat the idea of a long sentence without using one of the words found in the original. One sentence of Item 4 in the Dearborn C will illustrate this point as well as show another difficulty to be discussed later. Opposites of the underlined words are to be written in the spaces left for that purpose.

"*Few* would note his . . . , but . . . his *absence.*" "Presence" and "many" are the words supplied by the scoring directions. Children undoubtedly use "lots," "loads," "heaps," etc. as colloquial substitutes for "many." "Attendance" is surely the school opposite for "absence." A child writing, "Few would note his attendance, but lots his absence" would give good evidence of his grasp of the ideas for which the words are symbols, yet to score this as right would be a departure from the instructions and would require a subjective evaluation.

Another scoring difficulty appears in the order of the words filled in by the testee. Suppose the above sentence to have been completed in this way: Few would note his *many*, but *presence* his absence. A purely mechanical following of directions would score this answer as wrong. Indeed, it would be hard for one scoring thousands of tests not to score it as wrong, for one gets to a point where he scores by the "feel" of an

answer and his blue pencil makes a cross before he has thoroughly grasped just what is the matter with it. The lack of sense in the above example offends him before he realizes what causes the offense. Yet the test is an opposite test, and it may well be that the exceptionally clever child, grasping the idea quickly and realizing that he is working against time, reads but the underlined words and quickly sets down their opposites in the first space he comes to. Subjective scoring is necessary here to avoid error. About half of the forty-two errors found in this item were caused by this subjective pull against direc-

TABLE IV.—ANALYSIS OF ERRORS FOUND IN FORTY-ONE DEARBORN A TESTS

Subjective			Objective		
I	II	III	I	II	III
1	3	6	4	0	0
2	14	2	8	2	0
3	3	2	9	1	0
5	2	1	13	3	8
6	3	4	14	0	0
7	1	1	16	18	5
10	7	5	17	2	5
11	18	2			
12	5	2			
15	10	2			
Totals.....	66	27		26	18

Total errors—137.

Column I—number of the item of the test.

Column II—number of tests raised by the check.

Column III—number of tests lowered by the check.

The division of "subjective" and "objective" refers to the way in which the test is scored.

TABLE V.—ERRORS FOUND IN ONE HUNDRED NINETY-SIX DEARBORN A TESTS; BOTH SCORE AND CHECK MADE BY SAME PERSON

I	II	III	IV
49	147	43 (1)	19 (1)
..	...	25 (2)	4 (2)
..	...	17 (3)	8 (3)
..	...	10 (4)	3 (4)
..	...	3 (5)	3 (6)
..	...	7 (6)	
..	...	1 (8)	
..	...	2 (9)	
..	...	1 (14)	
..	...	1 (17)	
Totals...49	147	110	37

tions, about one quarter were caused by lack of understanding of directions, and the other quarter seemed to be oversights.

The results of an analysis of forty-one Dearborn A Tests are shown in Table IV. The items are classified as objective and subjective according to the way in which they are scored. They are further subdivided to show underscoring and overscoring. There are one hundred thirty-seven errors in all; there are seventeen tests, hence, there is an average of about eight errors to a question. The objectively scored tests show an average of $6\frac{2}{7}$ errors per item. Twenty-three errors, the greatest number found in any single item appear in No. 16, The Clock Test, which is scored from a stencil.

The major causes of error seem to be:

(1) Differences in subjective estimate; (2) misunderstanding of instructions; (3) carelessness (these errors are few in number but serious in nature); (4) errors not easily classified but present in greater degree than is comfortable to think of.

The chief cause of error in subjectively scored tests is failure to recognize that a child of six is not an expert draughtsman. The chief cause of error in objectively scored tests is oversight. Credit is not so much refused as overlooked. That both sources of error can be combatted is shown by Table V. Here, both score and rescore were made by the same person, the rescore being made immediately after fresh instructions were given her.

It is hard to stress too much the importance of careful and reiterated instruction of scorers, instruction as differentiated from direction. The scorer must be taught as well as told, and he must be re-taught at fairly frequent intervals if accuracy is to be secured. Even the most careful instruction is not enough to avoid error in scoring, every test paper should be checked, if possible, by a person other than the one scoring originally. If this is not practicable, the same person can profitably check on his own work (see Table V). Scoring is monotonous work, even if one be interested in it, and error is bound to creep into work whose salient feature is monotony.

The following are the significant points disclosed by a review of five hundred thirty Dearborn A and C Tests:

1. Many errors in scoring were present.
2. There was enough constant error to affect the entire group, and to make conclusions based upon individual mental ages computed from unchecked test scores open to doubt.
3. There was persistent underscoring.

4. Objectivity of scoring is not sufficient insurance against error. Very nearly as many errors were found in tests scored objectively as in those scored subjectively.

5. There is no sure way to prevent error in scoring, though frequent and careful *instruction* of scorers is very helpful and is probably essential to accuracy in scoring.

6. A check on scoring, preferably by someone other than the original scorer, is essential to accuracy in scoring.

Intelligence testing is so susceptible to error before the results come into the laboratory that inaccuracy should be guarded against wherever possible. Rescoring all tests will nearly double the cost of scoring but may very well prevent the publication of startling seeming-facts which, upon examination, will be found to have their origin in the blue pencil of the scorer rather than in the record of the child they purport to represent.

PHOTOGRAPHIC MEASURES OF READING ABILITY

MILES A. TINKER

University of Minnesota

Interest concerning the nature of eye movements in reading originated with the pioneer work of Javal and his associates during the years 1879 to 1892. They demonstrated that in reading the eye executes a series of consecutive movements interrupted by pauses. These pauses are comparatively long in duration and the movements are very rapid. Various methods have been employed to investigate eye movements[12] (pp. 74–89). The method of recording eye movements by photographing the image of a beam of light reflected from the cornea was developed by Dodge about 1901. This technique has been found very useful in reading investigations and is the method most employed at present in eye movement studies.

Photographic records of eye movements yield five measures of reading performance: (1) Duration and extent of eye movements; (2) number of pauses per line; (3) number of regressions per line; (4) pause duration; and (5) perception time per line. The first has been discussed in another report.[13]* Only the last four measures will be considered here.

Reports of photographic studies on reading usually give pause frequency (number of fixational pauses), regression frequency, and pause duration. A few investigations have also included perception time as one of the measures of reading performance. The purpose of this study was twofold: (1) To show the relative validity of the various photographic measures of reading performance. (2) To demonstrate how variation in the kind of material read affects the consistency of these measures.†

Apparatus and Method

A modified form of the Dodge apparatus for photographing eye movements was employed to record eye movements made during the reading of various kinds of material. This apparatus has been fully described by Tinker[12] and by Miles and Shen.[11] A beam of

* The approximate location of the fixations of the eye along the line of print can be plotted from the photographic records of reading. See references [2] and [12].

† The data for this study were collected at Stanford University in the spring of 1927.

light was directed into the subject's eye. On striking the eye the light was reflected from the surface of the cornea into a camera. Here a continuously moving film recorded changes in the direction of the beam of light which occurred with each movement of the reader's eye. The beam of light was interrupted at regular intervals of $\frac{1}{50}$ second. Hence the record appeared on the film as a series of dotted lines. Each line represented a fixational pause and each dot in the line had a time value of $\frac{1}{50}$ second. The reading of these records yielded the measures of reading performance cited above.

Four kinds of material were read before the camera: (1) A paragraph of scientific prose thirteen lines long. (2) Two paragraphs of algebra which were read separately. One consisted of the statement and solution of a problem and the other was descriptive material. They were eight and nine lines long respectively and both contained some formulæ. (3) A line of simple algebraic formulæ. (4) A line of complex algebraic formulæ.

Preliminary practice trials with closely related reading material were given to all subjects in order to accustom them to the apparatus and the kind of material read. Comprehension questions were asked after each reading. In a preliminary test the university students were timed with a stop watch while they read a news item. Sixteen university students and ten high school seniors were subjects in the experiment. All subjects read the scientific prose. The sixteen university students read the algebra and fifteen read the formulæ. All university subjects had passed the equivalent of two or more courses in high school mathematics. Visual acuity tests showed that all but one observer had emmetropic vision. This student was near-sighted but had been advised by an oculist not to use glasses for reading.

Results

The alternation of movements and pauses are the two most obvious phases of oculomotor behavior during reading. These rapid, sweeping movements which occur in reading as well as in moving from one fixation to another in any field of regard are called saccadic because of their nature. Erdmann and Dodge[5] were the first to conclude that perception in reading occurs during the pauses and that the unclear visual sensations arising during eye movement are not attended to. In later studies[3,4] Dodge showed conclusively that there is no clear vision during saccadic eye movements. Therefore in reading, pause duration is perception time.

Perception time is the most valid photographic measure of speed of reading.* Eye movement duration is so short that it is comparatively insignificant and is always neglected in computing photographic measures of speed of reading. Tinker[13] has shown that on the average about six per cent of reading time is consumed by eye movements. In no type of material did the movements take as much as eight per cent of the reading time. Furthermore it is very difficult or impossible to determine eye movement duration accurately from the ordinary reading record.† The average perception time per line is the product of the number of fixations and the average pause duration. Thus perception time includes both pause frequency and pause duration.

TABLE I.—COEFFICIENTS OF CORRELATION BETWEEN THE VARIOUS PHOTOGRAPHIC MEASURES IN PERFORMANCE IN READING EASY PROSE (PRODUCT MOMENT *r*)*

		Number of fixations	Pause duration	Number of regressions
Elementary school: seventy-two cases.	Perception time	.87 ± .02	.79 ± .03	.61 ± .05
	Number of fixations	.....	.33 ± .07	.74 ± .04
	Pause duration	.....	.....	.19 ± .08
High school and university: sixty-seven cases.	Perception time	.81 ± .03	.73 ± .04	.58 ± .05
	Number of fixations	.....	.16 ± .08	.70 ± .04
	Pause duration	.....	.....	.86 ± .02
Elementary school through university: one hundred thirty-nine cases.	Perception time	.86 ± .02	.88 ± .01	.82 ± .02
	Number of fixations	.....	.33 ± .05	.83 ± .02
	Pause duration	.....	.....	.61 ± .04

* Calculated from Buswell's data[1] (pp. 14–17).

* Throughout this report when we speak of reading ability we mean speed of reading. It is well to remember that photographic records show certain characteristics such as regressions and variations in pause duration which indicate comprehension difficulties. However, such factors always lengthen or shorten perception time and thus affect speed of reading. In reading investigations where a photographic technique is used comprehension is usually checked by questioning.

† Units of $\frac{1}{25}$ or $\frac{1}{50}$ second are the intervals ordinarily employed for interrupting the beam of light directed into the reader's eye. Intervals of at least $\frac{1}{100}$ second or shorter are necessary to insure fair accuracy in determining eye movement duration in moving from one fixational pause to another during reading. Where the successive pauses in reading are very close together (spacially) it is practically impossible to determine directly the temporal duration of the movements between them.

The validity of other photographic measures of reading ability can be obtained by comparing them with perception time. Table I gives the intercorrelations between various photographic measures of reading performance. The calculations were made from Buswell's [1] data. The subjects were divided into three groups as follows: (1) seventy-two elementary school pupils from Grade III to Grade VII; (2) sixty-seven students extending from high school freshmen to college; (3) Groups one and two combined. Since pause duration is quite stable after Grade VI, but is not in the early elementary grades, we have considered the elementary school group and the higher group separately as well as together. All read the same selections.

In our consideration of the results in Table I we should bear in mind that comparatively easy narrative prose was read. The number of fixations per line correlates highly with perception time in all groups: .87 in the reading of elementary school pupils; .81 in that of the high school and college group; .86 in that of the total group. This demonstrates that number of fixations per line is a very good measure of performance in reading easy prose. Exceptions to this generalization will be noted later. For the two sub-groups pause duration correlated with perception time .79 and .73 but for the total group the correlation was .88. A similar relationship is evident when number of regressions is correlated with perception time although the coefficients are somewhat lower: .61 and .58 for the sub-groups and .82 for the total group. These coefficients indicate that for group comparisons in this kind of reading (easy narrative) pause frequency is a very good measure of reading performance, pause duration a good measure and regression frequency a fair measure. In comparing a reading test performance with photographic measures of reading Gray[7] noted that the slowest readers made pauses of greater length and number than the most rapid readers. He also points out that regressions are a feature of slow reading but are largely avoided by fast readers.

One also notes in Table I that fixation frequency and pause duration show little correlation with each other in any group. However, regression frequency does correlate rather well with the number of fixations. Pause duration and regression frequency show very little relationship with each other in the elementary school group, a fair relationship in the total group, and a very good relationship in the high school and college group. This last high correlation (.86) is difficult to interpret.

Gray has described two types of rapid readers: (1) Those who reduce pause frequency to a minimum; and (2) those who reduce pause duration to a minimum. If these two types were common the relationships indicated by the above correlations would not hold when considering individual cases. In reading easy narrative, one does occasionally encounter a case of Type two but they probably never or very seldom occur in difficult reading or study. It has been shown[12] that in smooth, efficient reading, pause duration is the most constant measure of reading performance. The indications are that one is justified in employing photographic measures of reading ability in individual diagnoses. Freeman,[6] Gray,[8] and Hincks[9] have used this method to good advantage for studying individual cases with reading difficulties.

The question arises whether speed of reading easy narrative material corresponds closely to performance in reading scientific prose or study. It was found that speed of reading a news item (easy narrative) and perception time for reading scientific prose gave a correlation of only .46 ± .12. If we correlate performance in rapid reading and performance in study (calculated from Judd and Buswell's[10] data, Table IX) we obtain the following coefficients:

Comparison	ρ with PEρ
Perception time	.46 ± .11
Number of fixations per line	.43 ± .13
Pause duration	.43 ± .13
Regressions	.47 ± .11

These coefficients agree closely with our results. They range from .43 to .47 as compared with our value of .46. The coefficients indicate that there is only a small degree of relationship between performance in rapid easy reading and study or comparatively difficult reading.

It is also interesting to note whether the conditions shown in Table I hold for scientific prose or study. In Table II appear correlation coefficients between the various measures of reading performance. Two groups of results are given: (1) Those from the reading of scientific prose by our twenty-six subjects; (2) those from reading with directions to study. Twenty subjects were employed. The reading by both groups was done with a definite study attitude. Results from the two groups correspond remarkably well. Each measure of performance shows the same tendency in the two groups although the size of the coefficients vary somewhat. Pause frequency correlated highly with perception time in the two groups (.84 and .80). In both groups pause duration yielded a low, positive correlation with perception

TABLE II.—COEFFICIENTS OF CORRELATION BETWEEN THE VARIOUS PHOTOGRAPHIC MEASURES OF PERFORMANCE IN READING PROSE WITH A STUDY ATTITUDE (RANK CORRELATION)

Comparison	Our data: scientific prose (study attitude)	Calculated from Judd and Buswell's data: study attitude
Perception time *vs.* number of fixations	.84 ± .04	.80 ± .05
Perception time *vs.* pause duration	.18 ± .12	.28 ± .14
Perception time *vs.* number of regressions	.80 ± .05	.47 ± .11
Pause duration *vs.* number of fixations	−.27 ± .11	−.31 ± .14
Pause duration *vs.* number of regressions	−.15 ± .12	−.24 ± .14
Number of fixations *vs.* number of regressions	.80 ± .04	.67 ± .09

time (.18 and .28). In like manner regression frequency gave coefficients of .80 and .47 with perception time. In both groups pause duration showed a low negative correlation with pause frequency and regression frequency (−.15 to −.31). Regression frequency correlated well with pause frequency (.80 and .67).

A comparison of these results with those in Table I will indicate the influence of a change in reading attitude on certain photographic measures of reading performance. In comparing perception time *vs.* number of fixations, change of reading attitude had no effect. For study (Table II) the coefficients were .80 and .84 which are about the size of those in Table I. Change of reading attitude did cause a decided change for perception time *vs.* pause duration. The correlations are fairly high for reading easy narrative (Table I) but very low for study. Perception time *vs.* number of regressions showed a fair

TABLE III.—COEFFICIENTS OF CORRELATION BETWEEN PERFORMANCE IN READING SCIENTIFIC PROSE AND READING VARIOUS OTHER KINDS OF MATERIAL (RANK CORRELATION)

Performance compared	Scientific prose compared with:		
	Algebra	Simple formulæ	Complex formulæ
Perception time	.70 ± .09	.42 ± .15	.33 ± .16
Pause duration	.59 ± .11	.53 ± .13	.26 ± .17
Number of fixations	.66 ± .10	.78 ± .07	.53 ± .13
Number of regressions	.86 ± .03	.55 ± .12	.35 ± .15

relationship in both kinds of reading. In reading easy narrative pause duration *vs.* pause frequency gave moderately low positive correlations; in study, moderately low negative coefficients. For the pause duration *vs.* regression frequency comparison the correlations were fairly high for easy narrative (except with the elementary school pupils where it was .19) but low and negative for study. Number of fixations *vs.* number of regressions yielded a fairly high correlation in both types of material. Thus all those comparisons which involved pause duration were strongly affected by the change in reading attitude.

Table III gives correlations between performance in reading scientific prose and reading various other materials. The first column of coefficients shows that performances in reading scientific prose correlate fairly well with those in reading algebra. The coefficients range from .59 to .86. Correlations between scientific prose and simple formulæ indicate fair relationship (.42 to .78). However, comparison of reading complex formulæ with scientific prose shows only slight correlations (.26 to .53). The correlations given in Table III indicate that individuals are not consistent in reading performance when the reading situation is changed. The greater the change in reading attitude (brought about by variation in kind of material read or change in instructions) the less the relationship between reading performance in the different situations. It has been shown[12] that the reading attitudes in prose, algebra and formulæ are quite different.

The small degree of relationship existing between performance in reading easy narrative, scientific prose, and other types of reading indicates that the speed of reading one does not predict speed in the others. Investigations should be carried out to determine whether a like situation is present for comprehension in reading.

Summary and Conclusions

1. Photographic measures of reading performance were investigated to determine: (1) The relative validity of such measures, and (2) to show how variation in kind of material read influences the consistency of these measures.

2. Intercorrelations were determined between perception time, pause frequency, regression frequency, and pause duration in reading easy narrative and in scientific prose. Also correlations between reading performance in various kinds of material were calculated.

3. Perception time is the most valid photographic measure of speed of reading.

4. In reading easy narrative, pause frequency is a very good measure, pause duration a good measure, and regression frequency a fair measure of reading performance.

5. In scientific prose (study attitude) pause frequency is a very good measure, regression frequency a fair measure and pause duration a very poor measure of reading performance.

6. Only a moderate sized correlation (about .46) exists between performance in reading easy narrative and in reading scientific prose (study attitude). Reading performance in scientific prose yields a good correlation with performance in algebra, a fair correlation with performance in simple formulæ and a slight correlation with performance in complex formulæ.

7. Changes in the reading situation favor inconsistency in reading performance: the greater the change in reading attitude the less the correlation between reading performance in the different situations.

References

1. Buswell, G. T.: "Fundamental Reading Habits: a Study of Their Development." Supplementary Educational Monographs No. 21, 1922.
2. Dearborn, W. F.: "The Psychology of Reading." New York, 1906; also as Columbia Univ. Contributions to Philosophy and Psychology, Vol. XIV, No. 2, 1906.
3. Dodge, R.: Visual Perception during Eye Movement. *Psychological Review*, 1900, Vol. VII, pp. 454–465.
4. Dodge, R.: The Illusion of Clear Vision during Eye Movement. *Psychological Bulletin*, 1905, Vol. II, pp. 193–199.
5. Erdmann, B. and Dodge, R.: "Psychologische Untersuchungen über das Lesen auf experimenteller." Grundlage, Halle., 1898.
6. Freeman, F. N.: Clinical Study as a Method in Experimental Education. *Journal Applied Psychology*, 1920, Vol. IV, pp. 126–141.
7. Gray, C. T.: "Types of Reading Ability as Exhibited through Tests and Laboratory Experiments." Supplementary Educational Monographs No. 5, 1917.
8. Gray, W. S.: "Remedial Cases in Reading: Their Diagnosis and Treatment." Supplementary Educational Monographs No. 22, 1922.
9. Hincks, E. M.: Disability in Reading and Its Relation to Personality. Harvard Monographs in Education No. 2, Vol. II, 1926.
10. Judd, C. H. and Buswell, G. T.: "Silent Reading; a Study of the Various Types." Supplementary Educational Monographs No. 23, 1922.
11. Miles, W. R. and Shen, E.: Photographic Recording of Eye Movements in the Reading of Chinese in Vertical and Horizontal Axes; Method and Preliminary Results. *Journal Experimental Psychology*, 1925, Vol. VIII, pp. 344–362.
12. Tinker, M. A.: "A Photographic Study of Eye Movements in Reading Formulæ." Genetic Psychological Monographs No. 2, Vol. III, 1928.
13. Tinker, M. A.: Eye Movement Duration, Pause Duration, and Reading Time. *Psychological Review*. (In Press.)

FIRST IMPRESSION VERSUS SECOND THOUGHT IN TRUE-FALSE TESTS

MARY LOUISE LOWE

Ocean Park, California

AND

C. C. CRAWFORD

University of Southern California, Los Angeles

Problem.—It is frequently said that it is better to record the first answer that comes to the mind in taking a true-false test than to stop to meditate or reason on the question. This idea is based on the fact that true-false tests are in large measure recognition tests and that the first response to an idea previously before consciousness is sometimes more reliable than the responses that occur after mature reflection. The purpose of the present article is to shed light on the relative merits of first impression and second thought in taking true-false tests, using the scores on the tests as the criteria for evaluation.

Procedure.—Two types of procedure were used. One was the actual tabulation of specific changes of answer in true-false test papers answered under ordinary circumstances and without any thought of such an investigation on the part of the students involved. This procedure assumed that any answer that was not changed was a "first impression" answer, and that any answer that was changed was a "second thought" answer. This is not necessarily a safe assumption, since many unchanged answers are "second thought" answers which were not written down in both forms. The second type of procedure was designed to correct this weakness. It involved experimental control of conditions in a rotation experiment. A large class was divided into Groups A and B. A true-false test was prepared in two parts, Test I and II. Each test was mimeographed so that two spaces were allowed for answering each question. The first space was for "first impression" and the second was for "second thought." Group A was asked to take Test I by answering rapidly all questions in the "first impression" space and then to return and answer each with more mature reflection in the "second thought" space, reversing the previous decision wherever desired. At the same time, Group B was asked to take Test I by a "delayed answer" form of second thought. This consisted of reading all questions over without answering any of

them and then returning to answer each in the "second thought" space on the sheet. This was to prevent the record of the first impression from influencing the final second thought decision. After Test I was finished in this manner each group took Test II, but the methods were reversed, so that Group A employed the "delayed answer" procedure while Group B recorded both the first impression and the second thought. This rotation of methods and groups should eliminate any differences in the ability of the groups or difficulty of the questions. It should be said here that the students were all juniors, seniors, or graduates in courses in education.

Results.—Table I shows the result of the detailed analysis of test papers, tabulated according to total number of changes. Table II gives the same tabulated according to students, and Table III according to questions. The three tables are all in agreement in that they show a definite superiority of changes from wrong to right over changes from right to wrong, with the ratio being almost exactly two to one for the total number of changes.

The results of the rotation experiment are shown in Table IV. It is of interest to note that the "first impression" scores were significantly lower than the "second thought" scores, which fact is in agreement

TABLE I.—RESULTS OF DETAILED ANALYSIS OF TEST PAPERS, TABULATED ON THE BASIS OF NUMBER OF ANSWERS CHANGED

Number of answers	21,903
Number changed	641
Number changed correctly	426
Number changed incorrectly	215
Ratio of correct to incorrect changes	2 to 1

TABLE II.—RESULTS OF DETAILED ANALYSIS OF TEST PAPERS, TABULATED ON THE BASIS OF NUMBER OF STUDENTS WHO CHANGED ANSWERS

Number of students who changed one or more answers	262
Number who made only correct changes	128
Number who made only incorrect changes	42
Number who made some correct and some incorrect changes	92
Ratio of students making only correct to making only incorrect changes	3 to 1

TABLE III.—RESULTS OF DETAILED ANALYSIS OF TEST PAPERS, TABULATED ON THE BASIS OF NUMBER OF QUESTIONS FOR WHICH ANSWERS WERE CHANGED

Number of questions on which one or more answers were changed	294
Number of questions in which only correct changes were made	152
Number of questions in which only incorrect changes were made	54
Number in which both correct and incorrect changes were made	88
Ratio of questions having only correct to having only incorrect changes	2.8 to 1

TABLE IV.—RESULTS OF THE ROTATION EXPERIMENT COMPARING FIRST IMPRESSION, SECOND THOUGHT, AND DELAYED ANSWER PROCEDURE

Test	Method of answering	Group A (N = 50)			Group B (N = 39)		
		M	SD	SDM	M	SD	SDM
I	First impression	22.88	2.60	.26			
	Second thought	25.72	2.12	.30			
	Delayed answer	.....	....	...	22.92	3.24	.61
II	First impression	.....	....	...	19.56	3.07	.49
	Second thought	.....	....	...	20.90	2.95	.47
	Delayed answer	20.06	3.06	.43			

Summary

Method of answering	Sum of means	SDS
First impression	42.44	.61
Second thought	46.62	.55
Delayed answer	42.98	.74

Comparison

Methods compared	D	SDD	EC
First impression *vs.* second thought	−4.18	.82	−1.8
First impression *vs.* delayed answer	− .54	.90	− .2
Second thought *vs.* delayed answer	3.64	.92	1.4

NOTE.—Key to symbols:

N = number of students.
M = mean, or average, score.
SD = standard deviation.
SDM = standard deviation of the mean.
SDS = standard deviation of the sum of the means.
D = difference.
SDD = standard deviation of the difference.
EC = experimental coefficient, based on the formula,

$$EC = \frac{D}{2.78SDD}$$

with the findings shown in the previous tables. The "delayed answer" method gave practically the same result as the "first impression" method, and distinctly lower than the "second thought" method. This would indicate that the best procedure for a student to employ

in taking a true-false test would be to answer all questions and then return to revise the first answer rather than to read all questions through before answering any of them.

A very important factor that should be brought to attention here is the amount of changing of answers which took place when the students were asked to record their "second thought" decisions. There were two hundred forty-seven changes out of a total of 2414 answers, or almost exactly ten per cent. In other words, first impression decisions were still clung to in nine out of ten cases on more mature consideration.

Summary and Conclusions.—We may briefly summarize the findings of the study as follows:

1. Detailed analysis of voluntary changes of answers in 21,903 true-false test responses showed the correct changes to be almost exactly twice as frequent as the incorrect changes.

2. When the same changes of answers were tabulated by students it was found that a similar ratio resulted, showing that the predominance of correct changes was not due to the peculiar reactions of a few individuals.

3. When these changes of answers were tabulated by test questions it was found that the predominance of correct changes was not due to the peculiar nature of a few questions.

4. When students were asked to record their first impressions and later to record their second thoughts, it was found that the second thought scores were significantly higher.

5. There was no significant advantage shown in having each student read the questions over before answering them, since the "delayed answer" scores were almost exactly equal to the "first impression" scores.

6. Second thought is not different from first impression on nine out of ten questions, but the one case out of each ten is enough to make a significant improvement in scores.

THE HIGH SCHOOL STUDENT AND SCIENTIFIC METHOD

M. LOUISE NICHOLS

Department of Science, South Philadelphia High School

In December 1910, Dr. John Dewey in a presidential address to the Section of Education of the American Association for the Advancement of Science said: "Science has been taught too much as an accumulation of ready-made material with which students are to be made familiar and not enough as a method of thinking, an attitude of mind, after the pattern of which mental habits are to be transformed."

In December, 1926, Dr. A. J. Carlson in an address to the Section of Medical Sciences of the same association voiced a similar criticism of science as taught not only in medical schools but also in secondary schools and colleges. A teacher of medical students is presumably in a fairly favorable position for evaluating the results of training in science already received by high school and college graduates. Dr. Carlson expressed the strong conviction that the facts of science are even now over-emphasized to the detriment of scientific method and attitude.

In would seem that during the first quarter of the twentieth century not a great deal has been done in schools and colleges to get away from the practice of imparting systematized knowledge and towards the ideal of developing habitual dispositions of mind. Is it to be so in the next quarter of a century?

Satisfactory development of habits in general depends upon their early initiation and habits of mind are no exception to the rule. Time is lost if bad habits must be uprooted and good habits implanted in their stead. From this point of view early training in science is important, for scientific method thoroughly acquired may last a lifetime and may profitably be applied to the solution of many of life's problems. Without it stores of facts are of little value.

Scientific method, as all educated persons are nowadays aware, consists in the accurate observation and comparison of a number of particulars followed by the drawing of a conclusion, *i.e.*, making a more or less generalized statement from these particulars. Frequently, in order that the observations may be complete and comparable, voluntary alteration of the conditions is necessary. In this process ingenuity, not seldom of a mechanical kind, is required. The mental skills that the repeated use of scientific method might be

expected to develop are: First, accuracy and completeness of observation; second, facility in the perception of resemblances and differences; third, ingenuity in experimentation; fourth, ability to draw legitimate conclusions from given data.

If it be possible to foster habits of accurate comparison and observation, children given adequate opportunity for the practice of such skills ought to show some improvement at the end of a term or of a year. In order to test this supposition girls entering the school in September were given frequent opportunities to observe and compare concrete objects and materials, for example, flowers, seeds, seedlings, different kinds of soils, etc. Near the end of January, and therefore near the end of their first term, they were given a test of ability to perceive similarities in a number of diagrams illustrating forms of crystals. The diagrams were entirely new to them and they had not yet had any laboratory work on crystals. Early in the following February a class entering the school was given, prior to any instruction in science, the same test. A comparison is made below of the percentages of the two groups perceiving more and less than three resemblances, the range in each group being from no resemblances to six.

Comparison of Crystal Forms

Grade	Percentage More than 3	Percentage Less than 3	Number of Pupils
IXA, beginning term	13	63.5	200
IXA, ending term	31	44	200

On the same day in February that the test was given to the entering class it was given also to classes of higher grade.

Grade	Percentage More than 3	Percentage Less than 3	Number of Pupils
XB	62	16	50
XIA	73.8	5.6	35
XIB	72.6	3.3	30
XIIA	93	0	29
XIIB	95	0	21

It is evident that the figures show an increasing percentage of girls able satisfactorily to perform this task. Part of this improvement may be accounted for by the gradual elimination of the less able students. To allow for this a comparison may be made between a class of thirty-seven just entering IXA, academic course, and a class of twenty-one just entering XIIB, academic course. The sixteen lowest scores of the IXA group may be disregarded and the comparison made between the twenty-one girls best able to perform the task and the twenty-one girls of Grade XIIB.

COMPARISON OF CRYSTAL FORMS

Number of similarities	0	1	2	3	4	5	6	7	8	9	Total
Number of pupils, IXA	1	8	9	11	4	2	2	..	..	..	37
Number of pupils, IXA	..	..	2	11	4	2	2	..	..	..	21
Number of pupils, XIIB	..	..	..	1	5	5	7	1	1	1	21

Since in XIIA and XIIB there are none who perceive less than three similarities, the comparison between Grades IXA to XIIB may also be corrected by disregarding all those who fail to perceive as many as three and re-calculating the percentages of those who obtain scores higher than three.

GRADE	PERCENTAGE MORE THAN 3	NUMBER OF PUPILS
IXA	42	19
XB	74	42
XIA	79	33
XIB	75	29
XIIA	93	29
XIIB	95	21

An obvious criticism is that the improvement may be due not to repeated exercise in making comparisons in the laboratory but to increasing maturity. Evidence as to the relative importance of these factors may be obtained from the results of a test of ability to perceive similarities in diagrams of leaves given in February to students of Grade XA. All of these students had received one year's instruction in elementary science and were approximately of the same chronological and mental age. Some had received ninth grade instruction in the South Philadelphia High School, while some had entered Grade X from other schools. All of the papers were marked by a teacher unfamiliar with the work of the pupils in IXB and without knowledge of where their instruction had been received.

COMPARISON OF LEAVES

POINTS	0–3	4–6	7–10	NUMBER OF PUPILS
	PERCENTAGE			
South Philadelphia High School	33	52	15	111
Other schools	43	52	5	39

The extent to which a difference in intelligence influences the scores obtained in a test of this character is shown by a comparison of the scores of accelerated (mental age in advance of chronological age) and retarded pupils (mental age below chronological age).

CLASSIFICATION OF PARTS OF FLOWERS

NUMBER OF PRINCIPLES USED	0	1	2	3 OR MORE
	PERCENTAGE			
Accelerated Grade IX	45	5	50	0
Accelerated Grade X	15	10	60	15
Accelerated Grade XI	0	20	10	70
Retarded Grade IX	50	10	40	0
Retarded Grade X	22.5	25	42.5	10
Retarded Grade XI	5	15	40	40

Apparently it is possible by suitable instruction to develop in the majority of students a habit of looking for resemblances. Is it possible to improve the accuracy and completeness with which a series of observations is made? To get an answer to this question girls of Grade IXA were given exercises in observation. They were provided with tumblers of water on the surface of which was dropped some powdered dye. They were then asked to record everything that happened. A month later they were given another exercise, this time on the sequence of events in the burning of a match. The nature of the performance with the first of these tests may be illustrated by the scores of two typical classes, one of higher mental rating than the other. In the more intelligent class the maximum number of events correctly recorded was six, in the other the maximum number was five. Dividing the scores into higher and lower halves, the comparison is as follows:

DIFFUSION OF DYE

GRADE	PERCENTAGE PERCEIVING 4–6	NUMBER OF PUPILS
IXA, higher mental rating	28.3	21
IXA, lower mental rating	14	28

The maximum number of events observed in the burning of a match was ten. Dividing these scores also into a higher and a lower half the results from five classes, a total of one hundred twenty-four girls, are as follows:

BURNING OF MATCH

GRADE	PERCENTAGE PERCEIVING 6–10	NUMBER OF PUPILS
IXA	50	124

At the same time that the pupils of IXA were taking this test a number of girls in IXB were given a similar test on the burning of a candle. They, however, had not received special training in observa-

tion. The following May the girls of IXA, now IXB, were also given the test on the burning of the candle.

BURNING OF CANDLE

GRADE	PERCENTAGE 6–10	NUMBER OF PUPILS
IXB, without special training	37.9	124
IXB, with special training	67.7	196

The question arises whether the habit of close observation and comparison will continue to function when the objects to be compared are not present to the senses but have to be recalled and held as mental images while the comparison is made. Girls from the same groups, after having discussed the thermos bottle and the refrigerator but having made no comparisons were asked to write down the resemblances and differences of these two devices.

COMPARISON OF THERMOS BOTTLE AND REFRIGERATOR

GRADE	PERCENTAGE MORE THAN 3	NUMBER OF PUPILS
IXB, without special training	34	53
IXB, with special training	69.5	194

Can we test also the effect of training on the ability to generalize from particulars? In every well-conducted laboratory students are required to record the results of their observations and to draw therefrom valid conclusions. Will this practice bring about an increasing ability to generalize and to reason correctly? To test this sheets were given to each of the IXB science classes with the following data:

WORKER ANTS	TEMPERATURE, DEGREES F.	TIME, HOURS	BEHAVIOR OF ANTS
First lot	32	24	Motionless
Second lot	50	24	Move very slowly
Third lot	75	24	Very active
Fourth lot	96	24	Motionless

Write a conclusion.

Of ninety-one girls taking this test 39.5 per cent were incapable of drawing a satisfactory conclusion. They either repeated the particulars or wrote an entirely incorrect conclusion. When the girls returned to school the following September they were given a test of similar character with reference to the conditions necessary for the renewal of leaves on trees. The test was taken also by girls entering from other schools and by those taking the college preparatory course in their senior year.

GENERALIZATION, RENEWAL OF LEAVES

GRADE	PERCENTAGE SATISFACTORY	PERCENTAGE INCOMPLETE	PERCENTAGE FAILURE	NUMBER OF PUPILS
XA, South Philadelphia High School	30	32	38	121
XA, other schools	22	17	61	18
XIIA and B	21	71	7	14

With these results may be compared those obtained from a test concerning the velocity of transmission of sound.

GENERALIZATION, TRANSMISSION OF SOUND

GRADE	PERCENTAGE SATISFACTORY	PERCENTAGE INCOMPLETE	PERCENTAGE FAILURE	NUMBER OF PUPILS
XA	29	41	30	97
XIIA and B, college preparatory	33	47	20	15
XIIB, normal preparatory	68	20	12	25

The college preparatory class had had one year of introductory science in Grade IX and no further instruction in science until Grade XII. The normal preparatory class had had three and one-half years of science and were entering the second term of their fourth year. There is rather a close correspondence between the performance of the college preparatory students and those about to enter their second year (Grade XA). If allowance is made for elimination of the less able students the correspondence is still closer. On the other hand the three and one-half years of science taken by the normal preparatory students has apparently fostered in them the ability to generalize correctly from particulars. This is corroborated by a comparison of the scores of a XIIB normal preparatory class on the test concerning ants and temperature and those made by girls of more than average ability in IXB and XB.

GENERALIZATION, ANTS AND TEMPERATURE

GRADE	PERCENTAGE SATISFACTORY	PERCENTAGE INCOMPLETE	PERCENTAGE FAILURE	NUMBER OF PUPILS
IXB	36	31	33	60
XB	36	44	20	25
XIIB	74	26	0	19

A test requiring the application of simple, elementary principles learned in geography and science in the lower grades also brings out differences in the abilities of those with differing amounts of instruction in science. The students were asked to describe and explain everything that might happen in the sky, air and water if the sun shone on the sea from 6 A. M. to 6 P. M. on a summer's day. For each phenomenon correctly explained two points on the score were given. The

scores ranged from zero to ten, the highest score being obtained by only three students, one in each of the three highest grades.

EXPLANATION, SUN AND SEA

GRADE	PERCENTAGE MAKING 0	2–4	6–10	NUMBER OF PUPILS
IXB, superior group	43	57	..	30
XB, superior group	27	50	23	31
XIB	11	66	23	27
XIIA and B, normal preparatory	9	60	31	42
XIIA and B, college preparatory	28	52	20	15

Although the girls are of approximately equal average ability, the difference in the percentages of those not attempting an explanation or incapable of giving correct explanations is great. A large number of the failures were due to lengthy descriptions being given without apparent consciousness that the causes of the phenomena described had been entirely omitted. This may indicate that the tendency to seek an explanation becomes more marked the longer the girls have studied science.

The ability to perceive how apparatus should be arranged to accomplish a definite purpose bears no very definite relation to the IQ. Possibly it is somewhat increased by practice in the laboratory, but I am inclined to think in lesser degree than the abilities already discussed. In this connection it may be worth while to consider the results of a test given both to girls of Grade IXB and to those of XIIB. It consisted of a series of diagrams in each of which a lamp chimney, a candle and two blocks of wood were shown but differently arranged in each diagram. The task was to select the diagram in which the apparatus was best arranged to show the creation of a draft upward.

DRAFT UPWARD

GRADE	PERCENTAGE GIVING CORRECT ANSWER	NUMBER OF PUPILS
IXB, superior group	74	24
IXB, inferior group	78	33
XIIB, normal preparatory	85	20

Somewhat related to this ability is the ability to perceive spatial relations. In order to test this each student was shown diagrams representing vessels of the same size but of different dimensions, one being wide and shallow, the other high and deep. The following questions were then put: "If into each of the vessels is put one-hundred cc. of water will there at the end of twenty-four hours still be

one hundred cc. of water in each? Will they still contain equal amounts of water? Give reasons for your answers."

PERCEPTION OF SPATIAL RELATIONS

GRADE	PERCENTAGE MAKING POINTS 0–1	2–3	4–5	NUMBER OF PUPILS
IXA	31	57	12	69
IXB	12	51	37	246
XIB, normal preparatory	4	31	65	26
XIB, commercial	18	38	44	27

DISCUSSION

In the Twenty-seventh Yearbook of the National Society for the Study of Education, 1928, there appear reports of investigations concerned with the function of training in creating habits and attitudes. Chapter XIII, written by Dr. G. M. Whipple is entitled "The Transfer of Training." Among the papers summarized three are of particular interest in connection with the present study. The first, a monograph by Miss Hewins, deals with the training of powers of observation in high school freshman. "After dividing the pupils of several classes into equivalent groups on the basis of suitable preliminary tests, one group in each class was drilled on the observation of biological material. Final tests were given to the entire class on the observation of non-biological material. On the whole, the results justify the assertion that there are certainly conclusive evidences of transfer."

The second is a paper by Dr. H. O. Rugg on the experimental determination of mental discipline in school studies. "The results show that the practice received in the fifteen weeks' work in descriptive geometry did substantially increase the efficiency of 'mental manipulation' of both geometric and non-geometric elements, and that the degree of transfer, as would be expected, varied with the degree of similarity between the practice and the test activity. As for the agencies of transfer, these were found by Rugg to lie not only in the direct carrying over of specific habits, but also in better methods of attack, better attitudes of orientation, and an increased facility in manipulating many elements at the same time."

The third is a study by Dr. E. L. Thorndike also on mental discipline. He concludes that the causal factor most surely at work is "the intellect already existent. Those who have the most to begin with gain the most during the year." In his opinion the physical sciences are equal, if not superior, to languages and mathematics from any point of view.

After summarizing the series of investigations Dr. Whipple makes the statement that "it is not only probable, but fairly certain, that some of the most important agencies of transfer are to be found among the higher-level relations, in generalized attitudes, moods, ideals, sets, ways of going about mental operations generally."

Our results at the South Philadelphia High School seem to indicate that a certain proportion of the girls entering the school acquire little in the way of mental training. Probably for the reason that they have already reached their intellectual limit, they are incapable of acquiring generalized attitudes and ways of going about mental operations. Distinctively scientific method is impossible for them although by much repetition they can no doubt acquire a certain amount of useful information and rule of thumb method which may help them to lead more efficient and healthful lives. If the early formation of habits of accurate observation, comparison and reasoning is considered important for the remainder of the high school population it appears well worth while for them to take as many courses in science as possible, provided that the courses emphasize scientific method and attitude and are not principally concerned with transmitting information.

SPELLING BY RULE

CORNELIA R. TROWBRIDGE

George Washington High School, New York City

Most adults who are not born to spell correctly but have painfully mastered our fantastic English spelling have built up for themselves a working knowledge of rules and analogies. They may not be able to formulate the rules, but they know how plurals are formed, when a final silent *e* is dropped, when it is retained, when consonants are doubled, and how suffixes are added. As they reached the stage of mental growth when the mind generalizes, they constructed their own formulas. Through these generalizations they have, by natural stages, outgrown many of their difficulties.

Have pupils of high school age reached the stage of generalizing? Do they know that there are basic principles which will help them to spell correctly, or do they regard each word as a law unto itself, to be memorized as if it were a Chinese ideograph? To answer these questions a series of experiments was made recently in a large New York City high school. Its aim was to find how far the pupils realized the existence of established laws of word formation in English and could apply these laws. It sought that insight into mental processes which is the source of all sound pedagogy. Because such gross errors as *comming, realy, widness* had been found in Regents papers, a test was made first of the senior class without warning or explanation. They were given the following carefully planned list of fifty words:

1. modifies
2. tragedies
3. tries
4. trolleys
5. chimneys
6. answers
7. confirms
8. asks
9. studying
10. modifying
11. hurrying
12. coming
13. writing
14. shining
15. arguing
16. excitement
17. careless
18. loneliness
19. likeness
20. arrangement
21. widely
22. planned
23. stopper
24. whizzing
25. occurred
26. omitting
27. limited
28. benefited
29. befitted
30. benefiting
31. regretted
32. coolly
33. occasionally
34. immediately
35. practically
36. carefully
37. accidentally
38. magazine
39. vegetable
40. original
41. official
42. magnificent
43. eligible
44. believe
45. yield
46. shriek
47. receive
48. weird
49. seize
50. siege

The words illustrate eleven fixed and simple principles. Numbers 1–8 are forms ending in *s*. The first three of this group add *s* to *y* after a consonant; the fourth and fifth add *s* to *y* after a vowel; the sixth and seventh illustrate the fact that *es* cannot be added to a consonant without adding a syllable: *ashes* is two syllables, so would *askes* be. The next seven words are participles formed from verbs ending in *y* or in *e*, the *y* to be kept, the *e* to be dropped without doubling of the preceding consonant. Numbers 16–21 are derivatives formed by suffixes beginning with a consonant added to silent *e*. Next come ten words, numbers 32–42, illustrating the doubling of the final consonant when preceded by a single vowel, both in monosyllables and in words of more than one syllable, provided the latter are accented on the last syllable—as *befit* is and *benefit* is not. The formation of adverbs by adding-*ly* is illustrated in numbers 32–37. Then follows a grouping of six words, numbers 38–43, in which hard *c* and *g* must be followed by *a*, *o*, or *u*, and soft *c* and *g* by *e*, *i*, or *y*. The remaining seven words, numbers 44–50, have the puzzling *ie* and *ei*. Eleven distinct principles are covered in this list.

As the test was on word formation, teachers were asked to make sure that the words were written with the right terminations, *-s*, *-ed*, *-ing*, *-ly*, etc. The students recognized that all the words were in their active vocabularies, and that it was wholly reasonable to expect mastery of them in the senior term. They were not told that the words were in groups, each governed by a rule,[1] for one of the purposes of the experiment was to find whether they were in the habit of analyzing spelling difficulties and applying general principles.

The results were amply sufficient to prove that these high school seniors did not know how to generalize in spelling. Out of two hundred seventy-nine papers, seven were one hundred per cent correct and two more had only such errors as *conferms* and *benifited*, not due to ignorance of rules. Two hundred seventy pupils failed to apply clearly defined and simple principles, the number of their errors ranging from one to fifteen.

The results were further analyzed as to the number of rules violated by each student:

---

[1] *Rule* is used in this article to mean a principle of word construction which can be formulated.

TABLE I

NUMBER OF RULES VIOLATED	NUMBER OF PAPERS
0	9
1	36
2	70
3	67
4	50
5	22
6	19
7	2
8	3
9	1
	279

Over half the students wrote *cooly*. That to drop the *ly* left only *coo* meant nothing to them, though they had doubtless been taught that adverbs come from adjectives. The same pupil would write *modifies* and *trys*, *comeing*, *writting* and *shining*, *benefited* and *benefitting*, *carless* and *likeness*. They did not know that *askes*, *widly*, *regreted*, *vegatable* indicated impossible pronunciations. Spelling to them was a subject for study in the elementary grades and their attitude of mind toward it was still childish. Some of them were in utter despair of ever learning. Some spelled wholly at random. The test brought home to them their own shortcomings, and the possibility of mastering completely large groups of everyday words on their own responsibility as seniors. They were held to account for their shortcomings, but to go into detail concerning the method used for giving them the training needed, is not the purpose of this paper.

The same test given to the seventh term English classes produced very similar results and with the figures from the two terms as cogent arguments for its necessity, an attempt was made to bring home to all students the existence of rules in English spelling and their value. The teachers were asked to cover the eleven rules on which the senior test was based which were given to all pupils on a printed syllabus,[1] and to

[1] Outline of the rules from the school syllabus:

RULE		EXAMPLE
I	*Y* after a consonant plus *s*	*tries*
II	*Y* after a vowel plus *s*	*valleys*
III	*S* after a consonant with no added syllable	*asks*
IV	*Y* + *-ing*	*studying*
V	Silent *e* + *-ing*	*coming*

prepare their classes for a test. Of course, the human equation made for varying degrees of thoroughness in this drill. Some teachers were engrossed in other aspects of their work. Some did not believe in spelling rules. Some did not fully understand their application. Illness and absence interfered with carrying out some good intentions. But, after a month preparation, a test was given to all students of the second and third years. The first year pupils, who are in an annex, were not included in this study. The test was again a list of common words, exactly parallel with that given to the seniors, *i.e.*, fifty words arranged in groups illustrating the rules of the school syllabus. The results by terms were as follows:

TABLE II

TERM	NUMBER OF PAPERS	NUMBER PASSING 100 PER CENT	PERCENTAGE OF PERFECT PAPERS
III	680	19	2.8
IV	916	40	4.3
V	548	29	5.3
VI	489	59	12.
	2633	147	Average 5.6

This shows a gain over the record of the seniors. Where the latter had only 2.5 per cent of perfect papers, the rest of the school averaged 5.6 per cent. Moreover, the more mature the pupils were, the higher per cent was their response to teaching a reasonable basis for spelling.

To penetrate further their spelling consciousness, the papers of twelve classes were studied in detail, three sets each for the third, fourth, fifth, and sixth terms. They represent the results of the work

---

VI	Silent *e* + suffix beginning with a consonant	*wideness*
VII	Monosyllables doubling final consonants	*begged*
VIII	Polysyllables doubling final consonants	*regretted*
IX	Adverbs from adjectives by adding *-ly*	*coolly*
X	Vowel after soft and hard sounds of *c* and *g*	*innocent*
		magazine

XI *IE* and *EI*, taught by this jingle:

Put *i* before *e*
Except after *c*
Or when sounded like *a*
As in *neighbor* and *neigh*
And except *seize* and *seizure*
And also *leisure*,
Height, *weird* and *either*
Forfeit and *neither*.

of teachers whose interest and skill in teaching spelling varied. Classified as to the number of rules violated, the results were as follows:

TABLE III

NUMBER OF RULES VIOLATED	NUMBER OF PAPERS
0	52
1	83
2	79
3	80
4	47
5	32
6	14
7	3
8	2
	392

Comparison with Table I will show a marked gain in the power to apply rules.

A further analysis was made to find the relative difficulty of the rules, based on the number of pupils violating each one. Of the three hundred ninety-two pupils, thirteen per cent did not know how to add the termination *-s* for plurals of nouns or for verbs in the third person singular of the present tense. There were four per cent who did not realize that sound and spelling must correspond in such words as *modifying*, which none of them would pronounce *modifing* though scores wrote it so. To add *-ing* to a silent *e*, as in *writing* and *arguing*, is one of the commonest of word formations, yet 26.2 per cent, or over one-fourth, failed to get uniformly correct the four words grouped to illustrate it. This is certainly an amazing comment on how spelling is taught by rote instead of reason. Keeping final *e*, essential to right pronunciation, before suffixes beginning with a consonant, was little better observed. Here 21.9 per cent failed. There were not many errors in doubling final consonants in monosyllables. Only three per cent of the papers had errors in this group. But to know when to double the final consonant in words of more than one syllable was too difficult for 69.6 per cent. The difficulty is in placing the accent and here the untrained ear contributed to the high percentage of errors. Applying the next rule is most simple. Since adjectives + *ly* = adverbs, adverbs − *ly* = adjectives. But the percentage of errors was high. Over forty-one per cent did not know that an adjective must be left after *-ly* is cut off. To detect hard and soft *c* and *g* again requires a trained ear and twenty-eight per cent failed to recognize the

force of a most useful rule. For the *ie* and *ei* words, where application of the rule requires quick thinking, the percentage of errors was 42.8 per cent.

It must be remembered in judging these results that a perfect paper meant not only mastery of fifty words, but an understanding of basic principles, each applicable to dozens of words, and a power to analyze difficulties and trace analogies, which is the mark of a maturing mind. Moreover, we were in many cases asking our pupils to revise their mental processes, and that is no slight task. Pupils clung to their old habits. They would surreptitiously write down two spellings for words and see which "looked right." Like their older brothers and sisters, they would put in sequence *becomming*, *writing*, and *pursueing*, write together *benefited* and *benefitting*, all of which revealed their lack of logic. Strange inconsistencies of their papers showed convincingly how mechanical had been their memorization and writing of words.

It seems safe to draw certain inferences from the results of this test. First of all, that to pupils of high school age, rules can be taught. As Tables I and III show, of the eighth term pupils who took the test without previous drill, only 2.7 per cent observed all the rules. Of the less mature pupils in the second and third years who had been forewarned and in some measure prepared, 5.6 per cent applied them all successfully. Moreover, this method of teaching spelling is adapted to the stage of intellectual growth to be dealt with in secondary schools. As the pupils gained in maturity, they acquired increasing power to apply the rules. The analysis of the results in the sixth term and the third term classes whose papers were especially studied, showed that in the sixth only twenty-seven per cent failed on more than two rules, in the third sixty-three per cent failed on more than two.

In order to find out whether the students of other schools were in a similar state of ignorance, the same list of words was given in two other large city schools, through the cooperation of the heads of their English departments. One of them, a co-educational school, has a careful and constant drill on lists of words, but little practice on the rules. Out of 2191 pupils who took the test, 1.4 per cent passed with perfect papers. The system by which the results were tabulated showed the average number of mistakes per pupil in the various classes ranged from 2.5 among the seniors to 11.5 in the first term. The other school, a girls' high school, gives little definite instruction in spelling. It admits pupils only by special examination, maintains a very high scholastic

standing, and expects good spelling as a by-product. The classes tested were in the first, sixth, and eighth terms. Out of one hundred two papers, only two had no mistakes. The average number of errors was 4.7 in the first term; 5.5 in the sixth term; 3.6 in the eighth term. These results of a test of only a hundred pupils are of little statistical value but certainly indicate that even pupils of a high average mentality do not reason from the fixed principles of English word formation.

In the school first tested, where teaching spelling by rule has been tried out, it has been to most of the pupils a new approach to an old problem, which has produced interesting reactions. At first the students showed utter inability to see the application of rules and to draw analogies, to rely on themselves; but when they understood the method, it roused their interest because of its novelty, because of its sure results, because they were not dealing with the unrelated lists that make up the usual spelling lesson, and because the method was a challenge to reason quickly. It also proved a great encouragement to the disheartened and a stimulus to the careless. In the exercises prepared for the drills, absolute correctness was required. An entire exercise was rejected if the principle to be illustrated was violated in a single word, and to find that they could meet these rigid requirements was a surprise and a tonic. Pupil after pupil said with the *naiveté* of youth: "Those rules *do* help. I never could spell before." And once trained to be observant, they improved also in the spelling of words not coming under the rules taught.

The errors in spelling which we must correct in secondary schools date back to elementary-school days. They mean the failure of mere memorizing as a method of study. They mean that visualization has not given the right spelling. The fatal "initial error" has been committed and there are wrong habits to be overcome. The experiment tried has showed, the teachers believe, an effective way of dealing with a situation which it is folly to ignore. The students are beginning to see spelling as a constructive process, stimulating and fruitful. The teachers have learned where many difficulties lie and how this stage in education can be made "guided growth" in the mastery of an essential element of correct English.

THE UNIQUENESS OF "G"

C. SPEARMAN

University of London

If any event is more likely than another to quicken the progress of psychological mathematics, it is the entry on the scene of a mathematician so eminent and so free from prejudice as Professor E. B. Wilson. May I, then, venture to ask him for some further elucidation on certain very important points?

To begin with, however, I should like to quote with lively satisfaction his explicit statement that he has examined and found correct (except for misprints), not only the mathematics in my book, "*The Abilities of Man*," but also the various mathematical papers to which this book makes reference. A similar scrutiny of the mathematics in both book and references has now been carried out by the eminent British statistician Professor A. Bowley and, as he has been good enough to tell me, with a similarly favorable issue. This testimony on both sides of the Atlantic is all the more valuable, seeing that in past years these mathematics have had to endure groundless criticisms by persons less competent and more biassed.

With much pleasure, too, I note Wilson's agreement that the assumptions from which the mathematics take their rise have under the appropriate conditions been satisfied by actual observation (within the limits of the experimental error). But I would add that on this matter of observation, unlike that of mathematics, there can be no finality. In proportion as the experimental error becomes reduced in size, the appropriate conditions will have to be expressed with increasing precision and completeness. As regards some of these finer conditions, I would specially refer to my paper in the *British Journal of Psychology*, 1928, p. 99. At present, unfortunately, even authors who should know better continue to bring forward observation where even the grosser conditions (see "The Abilities of Man," Ch. X) have certainly not been satisfied.

Turning to Wilson's original contributions, he first and foremost demonstrates that the current attempts to determine "g" fail really to do so. And his strictures apply no whit less when the letter g is replaced by any other name claiming to denote ability of comprehensive scope and yet measurable by a single coefficient. Among such are, for example, "general intelligence," "IQ," "mental age,"

and so forth. Even Kelley remarks that all these names do in substance assume the existence of g.[1] Indeed, in default of any such assumption, the attempt to express in a single coefficient a person's powers of perceiving, of reasoning, of imagining, of attending, and so forth must needs be as futile as to set up a single coefficient for, say, a country's productivity in corn, in smokes, in music, and in brave men, or in any other incommensurable things.

In pursuance of the said purpose, Wilson starts from the fundamental observation of the following system of equations, often known as that of the tetrad differences.

$$r_{ap} \cdot r_{bq} - r_{aq} \cdot r_{bp} = 0 \qquad (1)$$

a, b, p, and q stand for every different set of four of the variables (here the abilities) concerned, whilst the r's indicate product moment correlational coefficients.

He then accepts my proof that from the system (1) there can be deduced the following system (2). He agrees—in contradiction to some erroneous statements made elsewhere —that this deduction has been done in a rigorous manner; that is to say, without assuming any particular frequency distributions of the values; without involving any approximations; and without any appeal to mere probabilities.

$$m_{ax} = r_{ag} \cdot g_x + (1 - r^2_{ag})^{\frac{1}{2}} \cdot s_{ax} \qquad (2)$$

where m_{ax} is the value of the variable a for the individual x; g_x and s_{ax} are two further values variable from individual to individual; r_{ag} is the correlation between m_{ax} and g_x. Furthermore, g_x remains constant for the individual x throughout all the variables concerned, a, b, c, . . . But s_{ax}, s_{bx}, s_{cx} . . . are uncorrelated both with each other and with g_x. The sigmas (taken over all the individuals) of the m's, the g, and the s's are in each case equal to unity. Means have been found to calculate from observation the values r_{ag}, r_{bg} . . . Unknown remain the g and the s's.[1]

Wilson proceeds to point out that, if the number of the individuals be denoted by k, the system (2) gives only k equations but $k + 1$ unknowns. Hence g remains within large limits undetermined; *it cannot be regarded as unique.*

Now, with all this I cordially agree. In fact, I was urging very much the same thing myself.[2] But then I cannot plead guilty when

---

[1] "Interpretation of Educational Measurements," 1927.

[2] See especially pp. xvii–xviii of the Appendix, "Abilities of Man."

Wilson rather reproaches me with not bringing forward actual instances of g being satisfactorily determined. For, by reason of what both of us have said, such instances have not been furnished by other workers in this field. And as for making the required determinations myself, actual measuring has so far not been my job; I have only tried to show others how to do it.

But now comes a serious point where the relations between Wilson and myself become more dubious, and I cordially invite his cooperation to clear things up. His writings have left on many readers the impression that the indeterminateness of g just mentioned is essential and final. Whereas I believed myself to show how it can be remedied. Since the trouble is that we have $k + 1$ unknowns and only k observed equations, why not make another observation to complete the toll? And for this purpose I suggested what amounts in substance to the following equation:

$$r_{ag} = 1 \tag{3}$$

where a is some one or other of the abilities.

If we can get this equation (3), then from it and (2) there at once ensues without a shadow of doubt the perfectly determinate value for g_x

$$g_x = m_{ax} \tag{4}$$

That, I submit, absolutely disposes of any objection whatever on this point from its *mathematical* aspect. The truth or not of (3) does not come before the court of mathematics at all, but that of observation. Has it been, or can it be, actually observed? I venture to reply that this has actually been achieved as nearly as could reasonably have been xpected. Unpublished work in our laboratory has more than once obtained for an r_{ag} values of 0.99. And although there would still appear to be needed some theoretical deductions from this value, on the other hand there are indications that nothing stands essentially in the way of raising it much higher still; in fact, as near as desired to unity.

The use of this discovery is two-fold. First, it meets any such theoretical objection as has been attributed to Wilson. For when once g has been determined (actually or even potentially) for any one of the set of variables at issue, then this determinate value holds good throughout all the others. Incidentally, this value of g_x furthermore serves to determine all the s's throughout all the set save only s_{ax}, which is anyway of negligible importance. The second utility of

(3) is not theoretical but practical. It supplies an ideal to which actual measurements of *g*, "general intelligence," "IQ" and-so-forth ought at any rate to approximate.

Another momentous contribution to the topic from Wilson has been what he calls the transformation of correlations. He supposes the original variables *a*, *b*, *c*, ... to be combined in varying proportions, each such combination being regarded as a new variable. He lets the total number of such new variables be the same as that of the old ones. Then he shows that this new set of variables can be constructed in an infinity of different ways that all satisfy the systems (1) and (2) differently, so that *g* becomes altogether indeterminate. Yet the new variables, he urges, "contain all the information of the old scores *m*, but the information is differently assembled." Hence, he argues, the satisfaction of (1) and (2) by the original variables *a*, *b*, *c*, . . . was not absolute but only "relative to the set-up" of these variables.

Almost all this, also, I can thankfully accept. But some of his readers have understood it in the sense of raising an essential objection to determining *g* as required for psychology. To any such conclusion I should strongly demur. For the said "set-up" of the original tests was—as already mentioned and as inevitable—far from being arbitrary; it had, instead, to accord with certain conditions. And the very first of these has always been that none of the abilities *a*, *b*, *c*, should "overlap" each other.[1] But by the transformation, this condition is at once violated. For instance, let two of the new variables be $(w_1a + w_2b)$ and $(w_3a + w_4b)$, where the *w*'s are the respective combining weights. Evidently, the two new variables "overlap" in respect of both *a* and *b*.

Many other difficulties have occurred to me in the way of applying these transformations; hence further developments of the theory from Wilson would be very welcome. I will only mention the following instance. He seems to allow his combining weights to take both positive and negative values; in other words, he admits both additions and

[1] Formulated more precisely, the doctrine of *g* enounces that on considering any test of ability and conceiving other tests in a series of increasing difference from the original one, then a point in the series can be found such that all the greater differences will satisfy the system (1) (without the correlations sinking to zero), whilst all the smaller differences will fail to do so. These latter smaller differences are said to be cases of overlap. Experience has shown it may be added —that in general such overlap does not occur until the resemblance is very obvious and striking.

subtractions of the original variables. Now, the additions are in psychology not unreasonable; different abilities and test-scores can in many cases be combined additively into a single more composite ability and score. But in what conceivable cases can one ability or score be reasonably *deducted* from another, one that does not overlap it?

But suppose that, to meet the preceding objection, all negative combining weights are excluded from the transformation. Then, I suggest, the transformation can no longer be effected in an infinity of ways satisfying (1) and (2) differently. In fact, I suspect that now the combination can only be effected in one single way. If this be so, then the indeterminateness of g disappears even from the sphere of transformations. And what, then, is the real significance, of the unique value obtained by the transformation. No one is more capable of telling us than Professor Wilson.

COMMENT ON PROFESSOR SPEARMAN'S NOTE

EDWIN B. WILSON

Harvard School of Public Health

Professor Spearman very courteously has sent me his contribution (The Uniqueness of "*g*") with the suggestion that I add something of my own to it. Before venturing to do so I must express my appreciation and remark that I take no exception to any of his statements, and I should further explain what are the only possible utilities of any comments I may make.

The applied mathematician may, if his temper be proper, be useful in three ways: First, when the postulates or primary assumptions of a branch of science have all been precisely formulated in mathematical equations he, by his greater manipulative skill, may develop new theorems in that science; and second, when the postulates have not been completely stated so that the problem is not yet mathematically definite, he may develop a number of alternative systems of deduction which may be useful in helping the scientist to make the system more definite by adding other postulates; and third, when postulates have been at least in some respects overstated, he may help in revealing duplication or inconsistency among the postulates. My review in *Science*, Vol. LXVII, March, 1928, pp. 244–248 and my contribution to the *Proceedings of the National Academy of Sciences*, Vol. XIV, March, 1928, pp. 283–291, were directed toward those ends. Whether they were, in fact, useful and whether such comments as I add below may be useful, others alone can judge.

Commenting on my injection of the transformation theory, Spearman writes new variables $w_1a + w_2b$ and $w_3a + w_4b$ and remarks that these evidently "overlap" with respect to both a and b. My notion was that all tests $a, b, c,$ had to overlap in respect to the common reliance of their scores upon the general intelligence, g, and that it was only in respect to the specific abilities s_a, s_b, that we tried to eliminate overlap. The proposition of overlapping would then scarcely be evident, but might be proved. We should write

$$a = r_{ag}g + \sqrt{1 - r^2_{ag}}s_a,\ b = r_{bg}g + \sqrt{1 - r^2_{bg}}s_b \tag{1}$$

$$a' = w_1a + w_2b = (w_1r_{ag} + w_2r_{bg})g + w_1\sqrt{1 - r^2_{ag}}s_a + w_2\sqrt{1 - r^2_{bg}}s_b \tag{2}$$

$$b' = w_3a + w_4b = (w_3r_{ag} + w_4r_{bg})g + w_3\sqrt{1 - r^2_{ag}}s_a + w_4\sqrt{1 - r^2_{bg}}s_b$$

with the proper conditions.

$$\sigma^2{}_{a'} = w_1{}^2 + w_2{}^2 + 2w_1w_2r_{ab} = 1,\ \sigma^2{}_{b'} = w_3{}^2 + w_4{}^2 + 2w_3w_4r_{ab} = 1 \quad (3)$$

to standardize the scales of a' and b'. Then we have

$$s_{a'} = w_1\sqrt{1 - r^2{}_{ag}}s_a + w_2\sqrt{1 - r^2{}_{bg}}s_b,$$
$$s_{b'} = w_3\sqrt{1 - r^2{}_{ag}}s_a + w_4\sqrt{1 - r^2{}_{bg}}s_b \quad (4)$$

and the correlation between $s_{a'}$ and $s_{b'}$ is

$$w_1w_3(1 - r^2{}_{ag}) + w_2w_4(1 - r^2{}_{bg}). \quad (5)$$

This can, in truth, not vanish unless one of the four values w is negative. Thus Spearman's suggestion of the postulate that no combining weights may be negative makes it impossible to take linear combinations of a and b and the suggestions of his last three paragraphs hang together mathematically.

I suspect, however, that there is some other postulate really implied in the proof just given, and I further suspect that this postulate is precisely one of those which I should want to have explicitly stated because of its presumed relation to the whole question of the determinateness of g, not in the sense of statistical determinateness, in which I judge Spearman feels that I did but illustrate his own position, but rather in the sense of that principle of relativity which I introduced along with the transformation theory.

To explain I must point out that among the equations used by him there are some which are quadratic, *viz.*:

$$r_{ab} = r_{ag}r_{bg},\ r_{ac} = r_{ag}r_{cg},\ r_{bc} = r_{bg}r_{cg}. \quad (6)$$

Now if we replace a and b by a' and b' as above and leave g to be determined, calling it g', we have similarly

$$r_{a'b'} = r_{a'g'}r_{b'g'},\ r_{a'c} = r_{a'g'}r_{cg'},\ r_{b'c} = r_{b'g'}r_{cg'}. \quad (6')$$

(I have left the third test c alone; we must have three tests to make any determination of g or g'; we never need any more than three except for increased determination, Note A, and if we have more than three, the tetrad equations must vanish statistically.)

Next from the definitions of a' and b' we have

$$r_{a'b'} = w_1w_3 + w_2w_4 + (w_1w_4 + w_2w_3)r_{ab} \quad (7)$$
$$r_{a'c} = w_1r_{ac} + w_2r_{bc},\ r_{b'c} = w_3r_{ac} + w_4r_{bc}$$

But

$$r_{cg} = \sqrt{\frac{r_{ac}r_{bc}}{r_{ab}}} \text{ and } r_{cg'} = \sqrt{\frac{r_{a'c}r_{b'c}}{r_{a'b'}}} \quad (8)$$

Is $r_{cg'}$ equal to r_{cg} as it must be if $g' = g$? Let us square and substitute for the primed expressions. Then

$$r_{ac}r_{bc}[w_1w_3 + w_2w_4 + (w_1w_4 + w_2w_3)r_{ab}] = r_{ab}[(w_1 r_{ac} + w_2r_{bc}) (w_3r_{ac} + w_4r_{bc})] \quad (9)$$

which reduces to

$$w_1w_3\frac{r_{ac}}{r_{ab}}(r_{bc} - r_{ab}r_{ac}) + w_2w_4\frac{r_{bc}}{r_{ab}}(r_{ac} - r_{ab}r_{bc}) = 0. \quad (10)$$

Unless the chosen values of w_1, w_2, w_3, w_4 satisfy this somewhat complicated equation the value of r_{cg} cannot equal that of $r_{cg'}$ and g' cannot equal g. Moreover as among Spearman's conditions are $r_{bc} - r_{ab}r_{ac} > 0$, $r_{ac} - r_{ab}r_{bc} > 0$, we see that to satisfy (10) one, at least, of the four values must be negative and thus that we must be permitted to subtract scores on the different tests if we are to combine them linearly so as to reproduce the original g, for any all-positive combinations we must get a different g.

What this all means analytically for the proof we gave at the outset is that we are not permitted logically to write two equations like (2) but must go over to

$$a' = r_{a'g'}g' + \sqrt{1 - r^2_{a'g'}}s_{a'}, \quad b' = r_{b'g'}g' + \sqrt{1 - r^2_{b'g'}}s_{b'} \quad (1')$$

with a new general g' replacing g, and new specifics $s_{a'}$ and $s_{b'}$ replacing the combinations of s_a and s_b. Moreover it may be stated that g' can be expressed linearly in terms of g, s_a, s_b, s_c; *i.e.*, the new general g' corresponding to a, b, c, involves the old general g and the old specific abilities s_a, s_b, s_c and the new specific abilities $s_{a'}$, $s_{b'}$, s_c involve the old general and the old specifics, it being impossible to assume without demonstration that the new specific s'_c for the unchanged test is the same as the old specific s_c of that test, and I should guess that usually these specifics would differ. All this really follows geometrically, though perhaps not obviously, from considerations of the orthocenter of a spherical triangle such as I gave in my condensed and rather difficult article in the *Proceedings* cited above.

Spearman's psychological postulate that we may not subtract scores prevents us from making those linear combinations of the tests which might conserve g and hence, relying on his postulate that g is definite we come to the theorem that we cannot add scores either. We may make no combination of tests. Is this a paradox? I do not know. And if it be a paradox how are we to get out from under?

Writing a year ago I had in mind two ways out: First, to permit subtraction as well as addition of scores; second, to admit frankly that g, insofar as it had yet been explicitly specified, was relative to the set-up. That the set-up was perfectly arbitrary I should never have thought; that on the contrary, every effort had been made to have each test overlap on every other only in respect of g and not at all in respect of any of the specific abilities I should take for granted; but that complete success should attend the effort is perhaps too much to hope and thus the possibility of combining tests analytically seemed to me worth considering, and as our effort was to eliminate overlap in the specific abilities wherever it might inadvertently have occurred it seemed to me not unnatural to permit combining weights to be negative. I do not understand that there is any serious difference here between Professor Spearman's position and mine—I am trying to grasp his psychological postulates and he is trying to see the possible significance of my mathematics. See Note B.

In the discussion above I have limited myself to three tests because three tests, whose intercorrelation coefficients satisfy certain inequalities, always give a solution for g. Consider now four tests. If the tetrad equations vanish we may, except for possible loss of statistical weight, throw away any one of the four, reduce the other three to g, s_a, s_b, s_c and thereupon separate the fourth test d into its parts $r_{dg}g + \sqrt{1 - r^2_{dg}}s_d$. If, however, the tetrad equations are not satisfied (within statistically appropriate limits) we have somehow to replace one of the tests by another or several of them by others so that the resulting system does not overlap (except with respect to the postulated common factor g). But how shall we do it? Mathematics cannot tell the psychologist how he shall do it. Mathematics can point out that if combining weights are allowed there are infinitely many ways of doing it by simple linear combinations of a, b, c, d, but that the resulting g will depend on the combining weights used. If combining weights are not allowed we may set up the problem of keeping tests a, b, and c and replacing d (psychologically, not algebraically) by one d' which shall satisfy the tetrad equations, but the solution for g thus obtained must usually differ from that which would be found if we kept the tests b, c, d and replaced a by a' in such a manner as to have the tetrad equations satisfied.

Thus, so far as I can see, the solution for g depends on what the psychologist determines it shall be. I do not know that either Professor Spearman or I would object to this. There is the possibility

that it throws some suspicion on the objectivity of g unless all psychologists agree to make the determination in the same way or in equivalent ways; we may have the general intelligence according to Spearman, that according to Kelley, another according to Thorndike, and so on—all different—merely because one worker considers that d overlaps in respect to s_d on s_a, s_b, s_c whereas another feels that a overlaps with respect to s_a upon s_b, s_c, s_d. This is perhaps inevitable; perhaps, however, the differences in the different determinations of g would not exceed the indeterminateness of any determination and would tend to vanish as improved determinations came to render all definitions more determinate. Frankly, this whole matter of overlapping or independence worries me. I seem to feel dimly that somehow in the last logical analysis the doctrine of g must imply that our specific abilities are distinct and discrete, like different chemical elements or different quantized states in physics and that overlap is inherently impossible but appears in our testing only because we have not had the finesse to set a battery of tests a, b, c, which use in addition to the general intelligence only certain separate special abilities of mutually exclusive groups of special abilities. This is possibly not different from Spearman's point of view as stated in his last footnote above.

And now I may go back to the author's suggestion that we set $r_{ag} = 1$ and be done with all sorts of indeterminateness. Fine! That would suit me provided it can be done psychologically rather than analytically. It is impossible that $r_{ag} = 1$ without making a the test of g. There is then no specific ability whatsoever, nothing but sheer general intelligence required to score in test a. Once we have come by any means whatsoever to an a which makes $r_{ag} = 1$ we may throw away all our scaffolding and identify m_{ax} with g_x, m_{ay} with g_y, etc. (As we approach the ideal condition, $r_{ag} = 1$, the specific part of a diminishes; when $r_{ag} = .99$, $\sqrt{1 - r^2_{ag}} = .14$ so that only fourteen per cent of specific ability remains and that may be considered insignificant as Spearman seems to imply; being no psychologist I can hardly judge. Incidentally I may express my satisfaction and admiration that a method has been found that will make r_{ag} so large at 0.99. I hope soon to see this work in print and shall look forward with particular interest to some discussion of the annoying question as to how much of the unreliability of this test a should be attributed to variability in the g's of the individuals at different times of mental testing, as their weights or pulse-rates might and would differ, and

how much of it should be attributed to still outstanding imperfections of the method of measurement.) Note C.

Thus we come out at the end with a result somewhat equivalent, as I understand it, to that with which Thorndike begins. He just plain defines the measure of g at the outset as given by scores on his CAVD (Completions-Arithmetic-Vocabulary-Directions). This at once clears the decks for action. Any other test b may be written as $b = r_{bg}g + \sqrt{1 - r^2_{bg}}s_b$ and we have at once the definition of the specific ability s_b, except perhaps for annoyances connected with distributing the unreliability among b, g, and s_b. A test c may be treated similarly and, if we can properly take account of the various unreliabilities, we can without further ado tell how much s_b and s_c as thus defined do, in fact, overlap. We do not need to worry about overlapping, we calculate it; we do not need to write tetrad equations and deal in higher trigonometry; everything is simple and straightforward procedure once we get to $r_{ag} = 1$. However, whether we proceed to this end by initial definition or by a complicated analysis may make a good deal of difference with our background and with our interpretation, and our choice of procedure may in fact have been inspired to no small extent by our point of view.

Note A.—The increased determination of g with the increase of the number of tests is a matter of some interest. The total determination may be written r_{tg} and the indetermination as r_{ig} with $r^2_{tg} + r^2_{ig} = 1$. Spearman discusses the team test t in the appendix of his book. If we introduce the notation

$$r_{ag} = \cos a,\ r_{bg} = \cos b,\ \cdots,\ Z = \cot^2 a + \cot^2 b + \cdots$$

$$t = \frac{a \cot a \csc a + b \cot b \csc b + \cdots}{\sqrt{Z(1 + Z)}}$$

$$r_{tg} = \sqrt{\frac{Z}{1 + Z}},\quad r_{ig} = \sqrt{1 - r^2_{tg}} = \frac{1}{\sqrt{1 + Z}}$$

If there are n tests the adjunction of an $(n + 1)st$ as good as the average of the n diminishes r_{ig} by less than $r_{ig}/(2n + 1)$ and it takes more than $3n$ new tests of equal average value to cut r_{ig} in two.

Note B.—Although it may be questioned whether there is any psychological significance in linear combinations of tests into teams, the analytical process of obtaining t suggests the mathematical problem of determining the teams u, v, which when taken with t will conserve g. As t contains all the information about g that can be had

from a, b, we must find $r_{ug} = r_{vg} = \cdots = 0$ and hence all the intercorrelations of t, u, v, must vanish. The resolution will then be

$$t_x = r_{tg}g_x + \sqrt{1 - r^2_{tg}}s_{tx},\ u_x = s_{ux},\ v_x = s_{vx1} \cdots$$

The analysis is involved but in case there are initially only three tests a, b, c we may find for u and v

$$u = u' \cos q + v' \sin q,\ v = -u' \sin q + v' \cos q$$

with

$$u = \frac{a \cot b \csc a - b \cot a \csc b}{\sqrt{\cot^2 a + \cot^2 b}}$$

$$v = \frac{(a \cot a \csc a + b \cot b \csc b) \cot c - c(\cot^2 a + \cot^2 b) \csc c}{\sqrt{(\cot^2 a + \cot^2 b + \cot^2 c)(\cot^2 a + \cot^2 b)}}$$

Note C.—It should be observed that the equation $r_{ag} = 1$, although it looks like a single equation, really is equivalent to as many as there are persons taking the tests just as the equation $x^2 + y^2 + z^2 = 0$ really forces $x = 0$ and $y = 0$ and $z = 0$.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY JAMES E. MENDENHALL

Research Associate, Institute of Educational Research, Teachers College
Columbia University

Intelligence Testing

A Comparative Study of Audito-vocal Digit Spans. Mary H. Young. The Psychological Clinic, Nov.-Dec., 1928, 170–183. A group of sixty-five children, Grades I and II, were tested in digit span forward, learning span forward, digit span forward, and learning span backward. Performances were compared.

The Iowa State College Reasoning Test. Thomas F. Vance. University of Iowa Studies in Psychology, 1928, No. 12, 72–77. Correlation between grade average in first quarter and (*a*) following directions was .33; (*b*) arithmetic, .32; (*c*) synonym and antonym, .36; (*d*) disarranged sentences, .43; (*e*) number series completion, .19; (*f*) analogies, .43; (*g*) information .38; (*h*) reasoning, .26.

The Larger Aspects of the Special Class. Fred Kuhlman. The Journal of Juvenile Research, Jan., 1929, 19–27. This study of 11,783 children revealed grade and age distributions of subnormals.

The Later Performance of Under-aged Children Admitted to School on the Basis of Mental Age. Edward A. Lincoln. Journal of Educational Research, Jan., 1929, 22–30. The writer finds that the admission of children to kindergarten or first grade on the basis of the Dearborn and Stanford Binet tests is warranted by the ratings of teachers and the later achievement of the pupils so admitted.

Measurement of Achievement

A Survey of English Grammar in the High Schools of Iowa, 1926–1927. George D. Stoddard. Journal of Educational Research, Jan., 1929, 13–21. About 7000 students, Grades IX to XII, were tested in February and in May in spelling, punctuation, errors in grammar, and common errors in language. Norms are presented.

Sex Differences in School Marks with Achievement Test Scores Constant. Theo. F. Lentz, Jr. School and Society, Jan. 12, 1929, 65–68. A study of one hundred eighty-eight girls and two hundred two boys in Grades II to VI showed a definite tendency for teachers to assign higher marks to girls than to boys when the criterion was score on Stanford Achievement.

Iowa Placement Examinations—a New Departure in Mental Measurement. George D. Stoddard. University of Iowa Studies in Psychology, 1928, No. XII, 92–101. Results tended to show that the "aptitude" tests predicted college grades as well as did the "training" test.

An Account of the Derivation and Standardization of the Virginia Biology Test. T. S. King. The Virginia Journal of Education, Dec., 1928, 161–166. The author reports methods of construction and norms.

Psychology of Learning and School Subjects

Studies in the Learning of English Expression. No. I. Punctuation. Percival M. Symonds and Baldwin Lee. Teachers College Record, Feb., 1929, 461–480. A total of six hundred sixteen compositions were analyzed for errors in punctuation. From Grades IV to XII graphs show an increase in usage, a decided decrease in omissions, and a practically constant number of errors. The frequency of omissions and of different kinds of errors in each of the several grades is presented.

Pupil Exercises Used in Elementary History Textbooks. Edward A. Nudelman. The Elementary School Journal, Jan., 1929, 351–357. Nine textbooks were analyzed. Great variations in types of exercises were found between texts. "What" questions, twelve per cent, and "Discuss" exercises, ten per cent, were prevalent.

The Purposes and Values of Oral Reading in the Intermediate and Upper Grades of the Elementary School. C. T. Gray. The Elementary School Journal, Jan., 1929, 335–343. The writer makes a case for oral reading, especially when the development of certain appreciational abilities rather than speed and comprehension are the desiderata.

The Voluntary and Independent Use of Reference Books in the Wheeling High School. Virginia K. Heinlein and Sophia H. Ervin. The School Review, Feb., 1929, 142–146. A study of three hundred boys and three hundred twenty-seven girls show an increasing use of reference books from the freshman to the senior year, especially in English and Literature, and in history classes. Of the reference books seventy-nine per cent of the students consulted the dictionary; fifty-one per cent, the encyclopedia; thirty per cent Who's Who, etc. Eighteen per cent reported that they did not understand the use of references.

A Source of Confusion in Spelling. Ernest Horn. Journal of Educational Research, Jan., 1929, 47–55. Through a compilation of possible methods of spelling phonetically the syllables of the word, "circumference," the writer estimates that there are about one trillion combinations.

Curriculum

Determination of a Content of the Course in Literature of a Suitable Difficulty for Junior and Senior High School Students. Mary Crowell Burch. Genetic Psychology Monographs, Sept., 1928. The writer tested over six hundred students in Grades VII to XII with selections from thirty-six most widely used books or passages from books by fifteen authors. Thereby, difficulty of content was determined.

Determining Basic Reading Materials through a Study of Children's Interests and Adult Judgment. Herbert B. Bruner. Teachers College Record, Jan., 1929, 285–309. A report of an excellent, intensive investigation in which the spontaneous questions of children, children's reaction to specialist's questions, adult judgments of prose selections, children's choice of materials—all were used to determine grade placement (interest and difficulty).

Determining Standards in English Composition. Roy Ivan Johnson. The School Review, (I) Dec., 1928, 757–767; (II) Jan., 1929, 44–48. A questionnaire study of the difficulties encountered by thirty-five leaders in conducting group discussions.

Objectives of United States History: the Aims of the Elementary School vs the Expectations of the High School. M. E. Herriott. The School Review, Feb., 1929, 107–116. With one hundred sixty-three elementary school and one hundred twenty-eight high school teachers regarding the important dates and personages in American history, there was little agreement between the two groups, or between either group and Washburne's lists.

A Core Vocabulary for Elementary School Pupils. C. G. Shambaugh and Olive L. Shambaugh. Journal of Educational Research, Jan., 1929, 39–46. A list of 1309 words common to all Grades IV to VIII was compiled on the basis of a free-association test. The word list is compared with those of Horn and Thorndike.

An Objective Basis for the Determination of the Objectives and Materials for a Course in Botany for Secondary Schools. Amos N. Merrill. Journal of Educational Research, Jan., 1929, 31–38. The writer compares the results of an intensive study of periodical literature (magazines and the like) with his analysis of the content of twelve widely used books in botany.

Child Psychology

Attendance at Moving Pictures as Related to Intelligence and Scholarship. Harold Ellis Jones. Parent-teacher, March, 1928. The investigator reports the following conclusions: (*a*) Approximately a third of each high school group attend the movies on an average of four times a month. (*b*) No differences were found in attendance between nine to eleven and twelve to fourteen year old children, once per week average for each. (*c*) No differences between urban and suburban groups. (*d*) Slight negative correlation between times attended and intelligence, the correlation being higher in the grades than high school.

Individual Differences in the Experiences of Children. Sarah E. Chase. Journal of Educational Method, (I) Nov., 1928, 62–73; (II) Dec., 1928, 136–146; (III) Jan., 1929, 194–198. An intensive study of the out-of-school activities of two hundred nine children in Grades IV, V and VI is reported. "Books read" and "movies seen" are further analyzed. The writer makes interesting comparisons between his study and those of others.

Social Adjustment through Kindergarten Training. Ada S. Woolfolk. Childhood Education, Jan., 1929, 264–268. A study of five hundred twenty-one children five to fourteen years old tested by the Family Welfare Society on the Terman Binet. This test showed that the group was inferior in IQ to Terman's unselected group. Comparison between kindergarten and non-kindergarten children of same superior original mental rating showed a slight superiority in mental age for school group. Similar results were obtained with retarded groups.

The Negro Child's Interest in Writing Poetry. Harvey C. Lehman and Paul A. Witty. Education, Feb., 1929, 346–354. A study of 5000 school children showed that a larger per cent of Negro than of white children reported having written poetry during a given week. The per cent for both groups decreased markedly with age.

Educational Psychology

A Comparative Study of the Performances of Stutterers and Normal Speakers in Mirror Tracing. Lee Edward Travis. Psychological Monographs, 1928, No. 12, 45–50. Comparing two groups (fifty each), the writer found in mirror tracing

that (*a*) right-handed stutterers are more facile with the left than with the right hand in comparison with the normal, (*b*) a greater per cent of left-eyed individuals among stutterers.

Some Notes on Orphan Twins. Robert A. Davis, Jr. The Journal of Juvenile Research, Jan., 1929, 28–34. With twenty-three orphan twins a correlation of .77 was secured by the Stanford Binet.

Methods of Research in Physical Education. C. H. McCloy. American Physical Education Review, Jan., 1929, 10–16. A survey of various methods concisely presented.

Sex Differences in Modern Languages as Measured by Standardized Tests. M. Eustace Broom. California Quarterly of Secondary Education, Jan., 1929, 125–128. With large numbers of cases measured in silent reading and vocabulary in both Spanish and French, "neither sex group has a clear advantage over the other in the skills measured by the three standardized tests."

Why Are More Boys than Girls Retarded in School? Ethel L. Cornell. The Elementary School Journal, (I) Oct., 1928, 96–195; (II) Nov., 1928, 213–226. A total of seven hundred forty-nine retarded pupils (five hundred twelve boys, two hundred thirty-seven girls) in Grades I and V were studied over a period of three years. This study showed the boys to be slightly younger chronologically, slightly older in mental age by Stanford Binet and by performance scale, slightly higher in IQ. On a personality rating scale boys were rated as more restless, more lacking in drive, more emotionally unstable, and more unsocial than girls.

Some Aspects to Be Considered in the Organization of a First Course in Psychology. University of Iowa Studies in Psychology, 1928, No. XII, 110–118. A group of one hundred eight students in a technical school were canvassed by a questionnaire. The group favored elaboration of the following: Knowledge concerning influences of heredity and environment; memory, attention, association, reasoning, thinking and conscious processes; knowledge of how we learn, laws of learning and their application; control of emotions; influence of personality in school and business.

Administration and Supervision

Overlapping of Courses in Education. Frederick E. Bolton. Educational Administration and Supervision, Dec., 1928, 610–623. The writer presents a tabulation of topics found in an analysis of five widely used textbooks in introduction to education.

Predicting Success in Practice-teaching. James H. Zant. Educational Administration and Supervision, Dec., 1928, 664–670. For two hundred records of teachers college students, the correlation between psychology and methods was .55; between psychology and practice-teaching, .30; between methods and practice-teaching, .32.

The Predictive Value of Three Specified Factors for Success in Practice-teaching. M. Eustace Broom. Educational Administration and Supervision, Jan., 1929, 25–29. With one hundred forty-eight students correlations were secured as follows: Average grade points per unit practice-teaching and (*a*) average grade points per unit in education-theory courses, .213; and (*b*) Thorndike equivalent score, .296; and (*c*) Thorndike trade decile rank, .044.

Impersonal Measurement of Teaching. F. A. Moss, Wm. Loman and Thelma Hunt. The Educational Record, Jan., 1929, 40–50. A test was given to 6667

students in twenty-eight land grant colleges at the end of the first year of chemistry. These results were secured: (*a*) reliability of test, .90; validity, .638; (*b*) range in school median scores, 37.25 to 87.0; (*c*) sectional medians: North Atlantic, 70; Southern, 59; Middle West, 56; Far West, 70; (*d*) those with high school chemistry exceeded average of those without by seventeen points; (*e*) large institutions (one lecturer), median, 46.62, (several instructors), 55.22; small institutions (one instructor), 48; (*f*) a greater achievement for smaller classes; (*g*) slight superiority for those taking course with Ph.D.'s rather than M.A.'s.

What Students in Certain Liberal Arts Colleges Take for the Bachelors Degree. Clarence Linton and F. B. O'Rear. Teachers College Record, Jan., 1929, 343–347. A study of six hundred forty graduates in thirteen colleges.

A Study of Teaching Loads in Junior High Schools in Wisconsin. S. P. Unzicker. The School Review, Feb., 1929, 136–141. A report of practices in twenty-three schools.

A Comparison of the Achievement of Superior Children in Segregated and Unsegregated First-grade Classes. The Elementary School Journal, Jan., 1929, 380–386. Two groups of pupils (*N*, 27 each) of equal intelligence served as cases. From Sept. to May, *under similar methods and materials of instruction*, the segregated group proved slightly superior on Gray Oral Reading (time and errors) and on Pressey Second Grade Reading, but inferior in Buckingham and in Iowa Spelling Scales.

The Status of the Sciences as Entrance Subjects in the College of Liberal Arts. H. F. Kilander. School Science and Mathematics, Feb., 1929, 170–178. Reports from two hundred ninety colleges indicate that approximately fifty-five per cent of the schools require science for entrance, more so in the West than in the East.

The Status of Hazing in American Colleges. Verner Martin Sims. School and Society, Feb. 9, 1929, 201–204. In fifty-six smaller colleges, hazing was reported as serious in zero per cent of the northern colleges, in fifty-six per cent of the southern, in twenty per cent of the western, and in fifteen per cent of the eastern. Eighty per cent of the colleges reporting a guidance or *laissez faire* rather than a prohibition policy do not have serious hazing. Only eighteen per cent of the colleges having hazing have student control, while fifty-five per cent of the colleges not having hazing have student control.

Educational and Vocational Guidance

A Pre-test for Courses in Vocational Guidance and Some Alleged Values of Its Use. Harvey C. Lehman and Paul A. Witty. The Vocational Guidance Magazine, Jan., 1929, 145–152. A check list of vocations from which subjects choose is presented.

Diagnostic Value of the Vocational Interest Test. Edward K. Strong. The Educational Record, Jan., 1929, 59–68. Of a group of one hundred fifty-six college seniors who filled out the Vocational Interest blank, forty-six per cent entered the occupation on which they scored highest in the test; twenty per cent, the occupation on which they scored next highest; and eleven per cent, the occupation on which they scored third highest.

Children Crying for the Moon. Agnes M. Conklin. The Journal of Educational Sociology, Jan., 1929, 263–277. A study of eight hundred one high school entrants regarding Terman Group (IQ average, 110), school marks, reason for

school attendance, high school intention (ninety-five per cent expect to graduate), ambitions for higher education (nominating college, sixty-three per cent), nationality, and language of parents.

Why High School Pupils Attend Summer School. L. R. Kilzer. The School Review, Feb., 1929, 132–135. Among four hundred twenty responses, 27.6 per cent report desire to complete four year course in less than four years; 19.7 per cent wish to make up points because of failure; 13.6 per cent want to make use of summer time, etc.

A Study of Job Shifting and a Proposed Remedy through Guidance. Robert Walker. The Vocational Guidance Magazine, Jan., 1929, 174–181. An analysis of school, employment, and interview records of two hundred fifty-two continuation school boys revealed that (*a*) thirty-five per cent left full-time school in Grade VI or below, thirty per cent at end of VII, twenty-eight per cent at end of VIII, ten per cent at end of IX; (*b*) job shifting occurred regardless of school training but decreased with the attainment of a higher grade level; (*c*) job shifting occurred regardless of intelligence level, but better adjustments to the job occurred among the more intelligent; (*c*) intelligence played no part in job selection.

College Athletics and Scholarship. Mark E. Hutchinson. School and Society, Feb. 2, 1929, 151–152. The returns show that athlete graduates (*N*, 40) in comparison with non-athlete graduates (*N*, 67) are superior in per cent obtaining degrees, in average grades, in intelligence score; spend more time in college; take a greater proportion of both hard and easy courses.

Measurement of Special Abilities

Stanford Motor Skills Unit. Robert Holmes Seashore. University of Iowa Studies in Psychology, 1928, No. 12, 51–66. The tests are described in detail. Correlations between these tests and typing, instrumental music, and competitive athletics (*N*, 50) averaged .25. Reliabilities of the motor tests, .75 to .94.

Relationship of the Intelligence Quotient and Scores on Mechanical Tests with Success in Industrial Subjects. Grayson N. Kefauver. The Vocational Guidance Magazine, Feb., 1929, 198–203. With one hundred one pupils in trade school, correlations were secured as follows: Between school course grades and (*a*) IQ (Terman group), .05 to .58; and (*b*) Stenquist, .07 to .65; and (*c*) MacQuarries, .15 to .63. Correlations between IQ and (*d*) Stenquist was .35, and (*e*) MacQuarrie, .28; between Stenquist and MacQuarrie, .42.

Five Studies of the Music Tests. Esther Allen Gaw. University of Iowa Studies in Psychology, 1928, No. XII, 145–156. A report of correlations between different music tests for children and for adults, and on different reproducing machines for the same test.

Correlation between Intelligence and Musical Talent among University Students. George Cutler Fracker and Virgie M. Howard. University of Iowa Studies in Psychology, 1928, No. XII, 157–161. With a group of two hundred thirty college students, the correlation between IQ and pitch was .32; and intensity, .01; and time, .13; and consonance, .09; and tonal memory, .10; and rhythm, .12.

A Measure of Art Talent. Norman C. Meier. University of Iowa Studies in Psychology, 1928, No. XII, 184–199. A test for æsthetic judgment of art is described; and averages for grade groups and for art and non-art groups are presented.

Miscellaneous

An Experiment in Character Measurement. John N. Washburne. The Journal of Juvenile Research, Jan., 1929, 1–18. An objective study of the choices of seventy-three children regarding a suppositious case, "an automobile, now or later" and two real cases, "candy, now or later" and "ten cents, now or later." If "later" was chosen, the reward was greater than for the "now" choice. Misbehavior was associated with the "now" choice.

Index Numbers in Educational Work. Harold F. Clark. Teachers College Record, Feb., 1929, 453–460. Index numbers are reported for cost of living for teachers, the price of school buildings, the price of school bonds, and the price of instructional supplies.

Expectation of Mental Disease. Horatio M. Pollock and Benjamin Malzberg. Mental Hygiene, Jan., 1929, 132–163. A study of 21,997 cases showed rates for various ages, sexes and nationalities.

Mentality and Delinquency. Enola B. Hamilton. Education, Feb., 1929, 331–345. Resume and conclusion of some earlier studies.

The Women Whose Names Appear in "American Men of Science" for 1927. Luella Cole Pressey. School and Society, Jan. 19, 1929, 96–1001. This study reports distribution by states, by colleges, by ages, by degree received, by interests, and by profession of six hundred eighty-seven women.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

A Textbook in Educational Psychology

Educational Psychology, an Objective Study, by Peter Sandiford. New York: Longmans, Green and Co., 1928. Pp. XIX + 406.

A reviewer should confront himself with the question: How satisfactorily has the author achieved the purposes which he has set for himself? Dr. Sandiford, in his Preface, states his main objective as follows: "In this book a serious attempt has been made to confine statements to the objective and verifiable data of carefully controlled observations and experiments . . . Since the method is comparatively new, considerable space has been devoted to the scientific foundations of the subject."

In the reviewer's judgment, Dr. Sandiford has accomplished his purpose in a most creditable manner.

The most conspicuous characteristics of the volume, as the reviewer sees them, are as follows:

1. The book gives a large amount of data from other sciences, especially from physiology, anatomy, biology, and biometrics.
2. It traces the historical development of most of the major topics.
3. It includes summaries of, and references to an unusually large number of scientific studies.
4. It contains the materials appraised as valuable with few, if any, omissions on account of difficulty.
5. It is written in a direct, lucid, succinct—almost summary-form.

These characteristics may be illustrated by the contents of the first chapter, "The Inheritance of Human Traits." It introduces the main facts and theories concerning biological heredity. The outstanding studies are listed or summarized, dated and placed in their historical setting. Some biographical data about Mendel and others are given. Although the researches of the Galton-Pearson type on human subjects are given most space, the important studies and

theories of Mendel and his followers, of T. H. Morgan and other exponents of the *gene* or *factorial* hypothesis, are presented in some detail, and all of the important statistical studies of twins, inaugurated by Thorndike, are summarized. The chapter contains many summaries in the outline form of brief statements or passages No. 1, 2, 3, etc. It includes a large number of technical terms such as those utilized by the Morgan school. It presents various types of technical formulæ and concepts such as simple and multiple correlations and standard errors of estimates. Yet, despite the intrinsic complexity of the material, the technical terms and concepts and the condensed form, the account is readable and intelligible as a result of exceptionally careful composition. The chapter, like all others, contains an excellent and lengthy bibliography for further reading, but no questions or exercises.

Dr. Sandiford's volume resembles, more than any other American works, those of Thorndike. In quantity and difficulty, it seems to lie between Thorndike's Briefer Course and the three volume series. In selection and organization of materials, it differs in some respects. Dr. Sandiford's volume gives relatively more space to physiology and anatomy and relatively less to instincts and fatigue. It includes chapters on topics which have largely developed since 1910. For example: "Intelligence: Its Nature and Measurement," "Conditioned Stimuli and Conditioned Reactions," "Educational Tests," and two chapters on "Improvement in Special Subjects" in which objective studies in the psychology of reading, spelling, writing and arithmetic are treated.

Although the volume is a comprehensive summary, it reveals a clear point of view which is consistently maintained. As regards method, the author renounces all interest in "unverifiable introspections and beliefs." Although he acknowledges respect for behaviorism, he means merely to express his faith in experimental methods. As regards explanatory systems in physiological fields, he favors those of Herrick and Child; in psychology, he harmonizes markedly with Thorndike. His agreement with Thorndike on many fundamental issues is apparent in his treatment of learning. Dr. Sandiford's own contributions to theory, which are frequent and significant, often consist in revealing the essential identity of conception in schemes which seem superficially to be different as in the case of the Thorndike-Judd theories of transfer; "associative shifting" and "conditioned reflex;" "trial and error" and "insight."

This volume, obviously the result of long, mature and sagacious deliberation and study, will be read with zest by all teachers of educational psychology. It will be appreciated, too, by mature and serious students of the subject. Although it resembles most closely other texts for a "second" or "advanced" course, it could probably be used by superior students in a comprehensive first course. A. I. GATES.

Teachers College, Columbia University.

THE NEW SCHOOL MOVEMENT

The Child-centered School, by Harold Rugg and Ann Shumaker, New York: World Book Co., 1928, Pp. XIV + 359.

The generation that is now passing off the stage of action has witnessed a vast number of educational changes, experiments and movements. And this educational unrest is but the reflection of a restless age. We live in a period of transition between two civilizations. Most of us who have reached middle life were born in a relatively simple agrarian society; today we find ourselves drawn ever more rapidly into a marvelously complex and highly integrated industrial society which is becoming more bewildering with every passing year and is unfolding before our eyes a world that has little in common with the world of our grandfathers. Educational change is consequently in harmony with the spirit of the times.

Among the many educational movements which characterize this period of social transition the rise of the so-called experimental or child-centered schools may well prove to be one of the most significant. Although the movement dates from the year 1896 when John Dewey founded his laboratory school in Chicago, we know altogether too little about it. This is due largely to the fact that it has not been adequately reported. To be sure, we have had numerous articles and a number of volumes from the pens of those who have been identified with the movement, but these have commonly been written with the bias of the uncritical enthusiast and advocate. The appearance of "The Child-centered School" by Rugg and Shumaker may therefore be regarded as an educational event of the first importance. Hereafter no member of the profession can plead the absence of a trustworthy account as an excuse for ignoring what the new schools are doing.

The book has many merits. It is extremely well-written, is interesting from beginning to end, and provides a critical appraisal of

the child-centered school. Although the authors are clearly in sympathy with the movement, they do not hesitate to point out what in their judgment are its shortcomings. The volume also is comprehensive. After introducing the new school the authors devote three chapters to the social and historical background. This is followed by two chapters of description and four of criticism. Chapters XI to XX inclusive, comprising approximately one-half of the book, are devoted to the creative arts. The account closes with one chapter on the "physical setting" and one on "critical retrospect." The appendix contains an excellent bibliography.

"The Child-centered School" will, no doubt, have its critics. Some will say that Mr. Rugg and Miss Shumaker have gone over bag and baggage to the new school movement, while others no doubt will contend that they have been too critical. Many perhaps will feel that the book is not well-balanced and that entirely too much space has been given to the creative arts. To such a criticism the authors would probably answer that they have deliberately emphasized this division of the program of the new school because of the relative failure of most schools and of American civilization to recognize adequately the rôle which the arts should play in life. But the query which the present reviewer would like to raise is of a somewhat more fundamental character. It seems to him that, in spite of the very fine sketch of the changing social order presented in the second chapter, the new school movement lacks the solid foundation of a carefully thought-out social theory. The authors set up *tolerant understanding* as one of the two major ideals of the school. While tolerant understanding is badly needed in our society, it can hardly be made to serve as a social philosophy. But whatever critics may say of the book, they cannot ignore the challenge to educational tradition which it contains. In "The Child-centered School" Mr. Rugg and Miss Shumaker have set a standard in both form and content which their colleagues will do well to emulate.

GEORGE S. COUNTS.
International Institute, Columbia University.

INTELLIGENCE TESTING

La Pratique des Tests Mentaux, by O. Decroly and R. Buyse. Paris: Alcan, 1928. Pp. XVI + 402.

This is a practical manual of intelligence testing for the psychologist and teacher. It is a survey of all types of intelligence tests, group and

individual, including a large number of French, English and American tests with a few from other countries. Few men are better equipped than Professor Decroly to make such a survey. He has personally studied the work of psychologists in this country and from the very beginning he has taken an active part in the testing movement in French-speaking countries. In 1906, just after the appearance of the first Binet Scale, he was at work trying it out and making suggestions for its modification. And ever since that time he has been trying new tests and making contributions of his own. The book does not attempt to describe the results and uses of intelligence tests, nor does it concern itself with the technique of test construction or the statistical manipuiation of results.

The Binet Scale is dealt with in detail and many of its revisions are described. A large chart shows the age-placement of the tests according to twenty-six different revisions of the Binet in several languages. The Vermeylen adaptation of the Rossolimo profile technique is described and will be of interest to American workers, especially to those who think that a general intelligence test should be made to yield a sort of picture of the different mental functions in addition to a total score.

The authors give as much space to individual performance tests as they do to individual tests of the Binet type, and this is very refreshing because the performance type of intelligence test is generally dealt with very briefly or totally misunderstood by most American writers. Decroly, however, was one of the first to protest against the predominantly linguistic nature of the Binet, and so it is natural that he should appreciate the value of the performance test and give it a proper emphasis in the field of testing.

About half of the book is devoted to group testing, giving a description of verbal and non-verbal tests. Many well-known American tests are described and complete French translations of the Terman Group Test and the Thurstone Test are given. Results for many different groups of students examined by the Thurstone Test are given. An English test by Ballard has also been translated and the results for Flemish and Walloon children are shown, the latter being at all ages superior to the former.

Among the non-verbal tests mentioned there is a translation of the Myers Mental Measure and also several new tests constructed by the authors. One of these is for young children from ages three to seven inclusive, and makes use of several ideas adapted from the Binet

Scale. Another has been constructed by Buyse for older children. In addition Decroly has a picture arrangement test and a picture cause and effect test arranged for group purposes.

The expense of the group test procedure and the necessity for economy is emphasized by the authors. Now the expense of the group test arises, according to the authors, in the provision of a separate printed booklet for each child, and, therefore, economy is to be achieved by tests that do not require such printed booklets. Hence the interest in the tests of Chapman and of Ballard for which no separate booklets are required, the child merely writing the answers on slips of paper. In this country we also have to fact the question of economy in group test procedure, but the test booklet is not considered a source of expense. Here the major expense consists in the time taken to score tests, and hence our efforts to simplify the scoring technique and so cut down the huge expense involved in scoring thousands of booklets. The cost of the booklet itself is a minor matter.

Decroly and Buyse have done a splendid piece of work in introducing many new tests to the psychologists and educators of French-speaking countries and there is much of interest to the American psychologist in their book.

R. PINTNER.

Teachers College, Columbia University.

A HANDBOOK ON TESTING APTITUDE

Aptitude Testing, by Clark L. Hull. New York; World Book Co., 1928. Pp. XIV + 535.

Hull has very properly taken the stand that intelligence tests, in view of their history and use, are essentially tests for the prediction of school aptitude. Hence his title is meant to indicate a general treatise on psychological tests, a treatise that is half hand-book and guide.

The book is divided into two parts, one on principles and one on methods. The first part contains an introductory chapter on the history of psychological tests. The second chapter concerns the tendency to Gaussian distribution of individual scores in measured traits and presents Hull's theory that the distribution of trait differences in the individual is also a normal Gaussian distribution, that, in other words, individuals tend to possess a moderate amount of most aptitudes and to be conspicuously high or low in relatively few.

A chapter of sixty-one pages describes the varieties of tests in use. Many plates and specimen sections of tests make this an interesting summary of test forms. Another chapter, equally interesting, is concerned with "anatomical and other alleged signs of aptitude," judgments from photographs, blond and brunette traits, Kretschmer's types, handwriting, pulse, and bloodpressure, and skull dimensions.

The next four chapters develop the mathematical theory of aptitude measurement. Tests are conceived as measurable samples of behavior selected for the prediction of behavior.

The problem of "general" and "specific" factors in the determination of aptitudes Hull meets with a group factor theory. "For practical purposes of aptitude prediction a rather large number of groups or constellations of aptitudes ultimately will be isolated." On this issue Hull's language, like Spearman's, might lead a naive reader to conceive these "factors" or "determiners" as concrete entities of a sort not meant by either writer. Hull's actual use of the notion of a "determiner" is as unobjectionable as Spearman's. It is a purely statistical abstraction which is made the theoretical basis for a very clear exposition of correlation, attenuation, the effects of weighting, the deduction of true relations from fallible data, spurious correlations, and the limitations of partial correlation.

The second part of the book is devoted to practical and detailed directions for testing. The analysis of the activity to be forecast, the assembly of a trial battery of tests, the administration of such a battery to a trial group, the selection of aptitude criteria, the final selection of a test battery, and the methods for combining test results for maximum forecasting efficiency are each covered by an adequate chapter.

Appendices give tables for converting ranks into linear scores, for $1 - r^2$ and for $\sqrt{1 - r^2}$, squares and square roots to 1000, and a multiplication table to 100×100.

E. R. GUTHRIE.
University of Washington.

Beauty and Truth

An Outline of Æsthetics, Edited by Philip N. Youtz. Vol. 1: *The World, the Arts and the Artist*, by Irwin Edman, pp. 88. Vol. 2: *The Judgment of Literature*, by Henry Wells, pp. 95. Vol. 3: *The Mirror of the Passing World*, by M. Cecil Allen, pp. 102. Vol. 4: *With Eyes of the Past*, by Henry Ladd, pp. 100. Vol. 5: *Scientific Method in Æsthetics*, by Thomas Munro, pp. 101. New York; W. W. Norton and Co., 1928.

De gustibus etc. is on the tip of everyone's tongue and in the back of everyone's dictionary. But who abides by the maxim? To become articulate on the canons of good taste and remain intelligible is as rare as the dodo bird. Most fortunate were the authors of "An Outline of Æsthetics" in the initial audience for which their work was planned. These five pleasant, unpretentious little books are the result of a series of lectures before the People's Institute in New York City. The People's Institute is precisely what the name implies. At one stroke, therefore, the authors sensibly cut away from the congested involutions of scientific nomenclature, and got down to readable, reasonable, enjoyable English. Like Huxley in the *Lay Sermons*, they had the task of explaining to the man in the street what it was all about—a task not so easy as it sounds. They (and Mr. Youtz in his lucid introductions) have come very close to carrying out their intention. The work is scholarly without being plethoric, clear without being thin. In spite of their own evident richness of background, the authors seldom talk down. On occasions, the language in which they couch their honesties rises to the poetic, effortless, and without a trace of affectation. Invariably they leave the impression that they themselves are sensitive to the best, and that they have given their best.

The series is introduced, in a sort, by Irwin Edman's *The World, the Arts and the Artist*, a most interesting attempt to show the arts in relation to each other through an underlying philosophy. In "Art and the Artist," the author sets forth the thesis that:

> To the extent that life has form it is an art, and to the extent that the established disorder of civilization has some coherence, it is a work of art . . . The breaking of a stick, the building of a hut, a skyscraper or a cathedral, the use of language for communication, the sowing or the harvesting of a crop, the nurture and education of children, the framing of a code of law or morals, the weaving of a garment, or the digging of a mine—all these are alike examples of art no less than the molding of a relief or the composition of a symphony.

In *Art and Civilization* he pursues the theme as applied to human organization:

> Æsthetic enjoyment and artistic creation are anticipations in our civilization of what the Good Life would really be . . . In a rationally ordered society, all work would have the quality of art, all enjoyment would have the immediate and glamorous character of æsthetic appreciation.

In his three subsequent chapters on the æsthetics of words, of line, color, surface, of tone, time and intensity, he dilates on his theme, rising to the fine generalization, in his *Epilogue*, that:

> The artist is the true revealer of the mystery . . . For in a work of art he has recognized one aspect of the One, clear and passionate and intense. Experience has become for him a moment of lucid flame; he is in the experience of beauty, to use the now well-worn language, for a moment at one with the One, at home in the Absolute.

Henry Wells follows with "The Judgment of Literature." There is an admirable humility in what he does not attempt to do. He sets down no hard and fast rules, he makes no dockets nor pigeon-holes. With a good deal of skill and discrimination he names the difficulties that must have beset him, too, and points out the tools and the angles from which those difficulties of taste and judgment may be attacked. If writers and readers impregnated and mellowed with sweet words and powerful words are at loggerheads on their impression, there is hope for the man in the street in his own shifting judgment. Mr Wells discusses fifteen grounds of critical dissension: Imitation and fancy, directness and convention, truth and idealization, the whole and the part, narrative and allusion, simplicity and virtuosity, clarity and obscurity, etc., etc., and comes to the conclusion that:

> I am moved to stress two virtues . . . Amid unceasing experiment we must strive above all for tolerance . . . Experience warns us that our own best ideals will neither be those of the preceding age nor of the age to come. Criticism today is after all merely the philosophical statement of changing taste. The first virtue is therefore modesty, and the second is like unto it: namely, silence . . . We may even have too much of art and beauty. We can very readily . . . have too much discussion of them . . . in the presence of true art the greatest wisdom is silence.

M. Cecil Allen's "Mirror of the Passing World," while most closely approximating a vade mecum for æsthetic judgment, left the present reviewer most dissatisfied—perhaps for that reason. The author follows the progress of development of painting, discriminates between

"subject" and "subject-matter," and makes mildly dogmatic pronouncements on the contribution of this nation and that to the world's æsthetic heritage. Addressed to putative artists, Mr. Allen's advice is too obvious; addressed to putative art lovers, it is too useless, since the gist of the last chapter seems to be: Get what you can wherever you can, but use what you find at home, in your own country.

Vol. 4, "With Eyes of the Past" by Henry Ladd, is wholly admirable. Perhaps its excellence lies in the nature of the material. It makes a wonderfully good companion piece for Wells's "Judgment of Literature," tracing, as it does, the change in standards of taste across time. The author's conclusions are posited on precise pronouncements by this and that child of the time. The insinuating undercurrent of irony is delectable and illuminating, not the least amusing of his innocent statements being his invitation to the ultra-moderns to become reactionary in order that they might best become revolutionary.

With Vol. 5, the last and the best of the series, this rapid and inadequate review comes to a close. "The Scientific Method in Æsthetics," by Thomas Munro. It is full of explosive and provocative material, rich in suggestion, and decidedly certain (the reviewer doesn't mean "cock-sure"!) in its setting forth of a reasoned, ordered method of procedure for finding out just what the æsthetic is, where it comes from, how to guide and harness it for human use, if it can be guided and harnessed. It binds form to flesh, so to speak, and indicates the positive relation that must exist, somewhere, between beauty and duty and energy and food and glandular secretion, temperature, environment, education—a practical, sensible, imaginative grasp of the problem such as, say, Cellini must have had, when he tackled the molding and heating and casting of the *Perseus*. Perhaps the fine edge of the flavor in this symposium would have been taken off, if the reviewer had read Dr. Munro's book first and Dr. Edman's last. He will chance it, and recommend the books henceforth to his friends in that order. But he will most certainly recommend them!

B. J. R. Stolper.

Lincoln School of Teachers College.